UNIVERSITY OF CALIFORNIA
Santa Barbara

Statistical Methods for Treatment Effects with Treatment Switching

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Statistics and Applied Probability

by

Fang-I Chu

Committee in Charge:

Professor Yuedong Wang, Chair

Professor David Hinkley

Professor John Hsu

Professor Wendy Meiring

December 2015

The Dissertation of
Fang-I Chu is approved:

Professor David Hinkley

Professor John Hsu

Professor Wendy Meiring

Professor Yuedong Wang, Committee Chairperson

November 2015

Statistical Methods for Treatment Effects with Treatment Switching

by

Fang-I Chu

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my advisor, Professor Yuedong Wang. His continuous guidance, support, and care make the research presented in this dissertation possible. I am grateful that Professor Yuedong Wang has inspired me with great ideas in my research work, giving me career advice and great help in many other areas. I also would like to thank Professor David Hinkley, my committee member and mentor through graduate study. I would not be able to get through confronted struggles without his consistent patience and caring guide. My gratitude also goes to Professor John Hsu and Wendy Meiring, my committee members, who encouraged me greatly, and gave precious advice about career and life.

I would never have been able to finish my dissertation without the guidance of the faculty members, help from friends, and support from my family. My time at UC Santa Barbara was made enjoyable in large part due to the pleasure to study and work with many great graduate students who have become my lifetime friends. I also want to thank them for their help and useful discussion during my graduate study.

Finally, I would like to thank my parents, Chun-Lan Chen and Jer-Hrong Chu, for all their unconditional love and support, allowing me to pursue my academic career with full devotion.

Curriculum Vitæ
Fang-I Chu

Education

2015	Doctor of Philosophy in Statistics and Applied Probability, Department of Statistics and Applied Probability, University of California, Santa Barbara
2011	Master of Art in Applied Statistics, University of California, Santa Barbara
2009	Master of Art in Mathematics, San Francisco State University
2006	Master of Art in Economics, San Francisco State University
2004	Bachelor of Art in International Business, I-Shou University, Taiwan

Experience

2015	Intern Statistician, Santa Barbara Actuaries Inc.
2011-2015	Computer Systems Assistant, Department of Statistics and Applied Probability, University of California, Santa Barbara.
2013	Intern Biostatistician, Shire plc, Wayne, PA.
2010-2015	Teaching Assistant, Department of Statistics and Applied Probability, University of California, Santa Barbara
2009-2010	Readership Assistant, Department of Statistics and Applied Probability, University of California, Santa Barbara.
2007-2009	Readership Assistant, Department of Mathematics, San Francisco State University
2007	On-line Teaching Assistant, Department of Mathematics, San Francisco State University

Abstract

Statistical Methods for Treatment Effects with Treatment Switching

Fang-I Chu

Treatment switching occurs frequently in clinical trials and other medical studies. Ignoring switching usually leads to biased and inefficient estimates of the treatment effect. More research is necessary to account for possible correlation between switching time and survival time. This thesis consists of two topics, treatment switching in clinical trials in Chapters 1-4 and treatment switching in dialysis studies in Chapters 5 and 6. Our research work addresses the treatment switching scenario through three approaches: (1) include a frailty term to accelerated failure time models, along with a logistic model to accommodate prognostic status, (2) evaluate treatment effect based on the latent event time and (3) quantify the switch effect using a varying coefficient Cox model. The proposed methods are applied to investigate the dynamic effect of change in dialysis vascular access in dialysis patients. The statistical analysis using above methods consistently suggests that hemodialysis (HD) patients with access change from catheter (CATH) to arteriovenous (AV) have lower mortality risk than those without access change and early switching is more beneficial.

Contents

List of Figures

List of Tables

Chapter 1

Treatment Switching in Clinical Trials

1.1 Introduction

Treatment switching occurs frequently in clinical trials (Morden, Lambert, Latimer, Abrams and Wailoo 2011). It happens when patients fail to comply with the original assigned regime. There are two situations: drop-in and drop-out. Drop-in occurs when disease progresses in patients of the control group such that these patients require treatment. Drop-out, on the other hand, happens when patients in treatment arm have difficulty to stick with the procedure, possibly due to side effects caused by the treatment or a change of will. Fail to account for either situation as we attempt to evaluate the treatment effect would lead to potentially biased conclusions.

We review some existing methods in Sections 1.2 to 1.5 and discuss limitations of the existing methods in Section 1.6.

1.2 Intention-To-Treat and Per-Protocal

The intention-to-treat (ITT) is the traditional method to cope with the treatment switching in clinical trials. It ignores the switching between treatment arms, and draws conclusion based on treatment arms which patients were randomized to. While the statistical analysis can be carried out easily, the ITT method may lead to underestimation of treatment effect due to the diluted treatment effect and consequently may lose power.

The Per-Protocol (PP) approach, on the other hand, involving analysis about patients according to the treatment they actually received, usually is used to supplement an ITT analysis. PP censors patients as they switch or excludes any switching patient (Fergusson, Aaron, Guyatt and Herbert 2002). By doing so, PP may introduce selection bias by not accounting for relation between the switching and the prognosis of disease. Another approach is to assess the effect of treatment actually received by a patient in the Cox proportional hazard model that includes treatment as a time-varying covariate. The model takes the form:

$$\lambda_i(t) = \lambda_0(t)\exp(\psi R_i(t)),$$

where ψ is the treatment effect, $\lambda_0(t)$ is the baseline hazard function and $R_i(t)$ takes a value of zero while the ith patient is receiving the control and 1 while they are receiving the experimental treatment. This approach can break the randomization balance and is subject to potential selection bias as the PP method, if switching is related to prognosis.

1.3 Adjusted Hazard Ratio Method

Law and Kaldor (1996) proposed an adjusted Cox model based on the observation that patients can be divided into four groups according to their switching patterns. Denote control and treatment arms as C and T. The groups CC and TT stand for the patients who complied with the schemes in control and treatment arms, while groups CT and TC represent those who switched from control to treatment and vice versa. They assumed that the hazard ratio, which stands for the relative effectiveness of control and treatment, is the same at the time of switching treatments as at the time of randomization. This assumption is questionable since the relative effectiveness of the control and treatment at the time of switching could be smaller, due to the patients who switched being at an advanced stage of disease when the treatments are less effective. Furthermore, the grouping criterion is based on a future event; that is to say, a patient randomized to control C will be in group CT if the patient switches to treatment regime later. The use

of covariates depending on future event is problematic for proportional hazards regression.

1.4 Accelerated Failure Time Model

1.4.1 Counterfactual Time

Robins and Tsiatis (1991) introduced an accelerated failure time model to estimate the true efficacy of a treatment. Consider a randomized trial with two arms, where $R_i = 0$ indicates patient i in the control arm, while $R_i = 1$ indicates patient i in the treatment arm. Let T_i be the actual observed survival time of the ith patient. A counterfactual event time U_i that would have been observed for a patient if he/she has not been switched to treatment arm is computed based on an acceleration factor e^{ψ}. The factor e^{ψ} stands for the magnitude of multiplicative effect on survival time from receiving treatment. Assuming treatment is beneficial, then $\psi < 0$. It is assumed that U_i is independent of R_i based on randomization. Denote the switching time as S_i and switching status as $\triangle_i$, where $\triangle_i = 0$ for no switch and 1 for switch. Switching is only allowed from the control to the treatment arm. Robins and Tsiatis (1991) and Branson and Whitehead (2002) made the follows assumption for the counterfactual time

$$U_i = \begin{cases} T_i, & R_i = 0 \text{ and } \triangle_i = 0, \\ S_i + e^{\psi}(T_i - S_i), & R_i = 0 \text{ and } \triangle_i = 1, \\ e^{\psi} T_i, & R_i = 1. \end{cases} \tag{1.1}$$

The causal AFT model relates counterfactual time U_i to observed survival time T_i. Maximum likelihood estimation for treatment effect is obtainable if one specifies a parametric distribution for U. However, the AFT model fails to account for the possible correlation between U_i and S_i and neglects the fact that disease prognosis might be strongly related to the change in treatment regime.

1.4.2 Iterative Parameter Estimation

Branson and Whitehead (2002) used an iterative parameter estimation (IPE) algorithm to compute randomization based effect estimators (RBEE) for ψ. T_i and S_i are assumed to be independent. The algorithm is motivated by the need to adjust the treatment estimate to account for patients in the control arm switching to the treatment arm. The method is implemented in the following steps:

Initiation Obtain an initial estimate of ψ, ψ_0, by comparing the randomized groups using a parametric AFT model.

Iteration At iteration j, transform the observed survival times for all patients who switched treatment in control arm to create projected counterfactual

times using

$$U_i^* = S_i + e^{\psi_{j-1}}(T_i - S_i),$$

where ψ_{j-1} is the estimate of ψ at iteration $j-1$. Fit an AFT using U_i^* to obtain an updated estimate ψ_j. Repeat the iteration step until $|e^{\psi_j} - e^{\psi_{j-1}}| < \epsilon$, where ϵ is a small positive number that controls the precision.

1.4.3 Estimation Using Joint Parametric Model

Let $C_1, \ldots, C_n$ be independent non-negative censoring times which are independent of observed survival times. Let T_i be the survival time for patient i with associated censoring time C_i and censoring indicator δ_i. The observations are $Y_i = \min\{T_i, C_i\}$, and $\delta_i = 1$ when $T_i \leq C_i$ and $\delta_i = 0$ otherwise. Let S_i be the switching time of patient i from the control to the treatment arm and the treatment history can be considered as a binary process $X_i(t)$ with 0 representing control and 1 representing treatment. For a patient who switched to the treatment arm, the counterfactual time $U_i = S_i + e^{\psi}(T_i - S_i)$ can be rewritten as

$$U_i(\psi) = \int_0^{T_i} \exp\left[\psi X_i(t)\right] dt.$$

Robins and Tsiatis (1991) assumed that U_i is identically distributed across the randomized groups. However, if T_i is a censored survival time, then $U_i(\psi)$ is censored at

$$D_i(\psi) = \int_0^{C_i} \exp[\psi X_i(t)]dt,$$

where $X_i(t)$ may potentially depend on factors such as underlying prognosis that causes change in treatment regime. Therefore, even when censoring on the T-scale is non-informative, censoring on U-scale is likely to be informative which is illustrated as follows. Assuming the treatment is beneficial, that is $\psi < 0$, we have

$$\begin{aligned}
P(U_i > C_i | R_i = 0, \triangle_i = 0) &= P(T_i > C_i), \\
P(U_i > C_i | R_i = 0, \triangle_i = 1) &= P(S_i + e^{\psi}(T_i - S_i) > C_i) \\
&< P(S_i + (T_i - S_i) > C_i) \\
&= P(T_i > C_i).
\end{aligned}$$

That is to say, individuals in the control group who switched to treatment are more likely to be censored than those did not switch since experimental treatment increases the counterfactual time U_i.

Walker, White and Babiker (2004) assumed a parametric model for U_i and S_i with joint density $f(s)f(u|s)$. The aim is to remove the informative censoring from the dependence between censoring time C_i and S_i. The problem of informative censoring is resolved whenever the censoring time on the U-scale is independent

of U_i conditional on the switching time S_i. Walker et al. (2004) proposed a parametric model that consists of three parts:

1. Causal model: AFT model defined in (1.1) that relates U_i and T_i.

2. A bivariate frailty model for U_i and S_i: this model is based on the assumption that U and S are independent given a common frailty. Denote $H_u(u)$ and $H_s(s)$ as the marginal cumulative hazards for U and S. Walker et al. (2004) considered two distributions for the common frailty: positive stable distribution and gamma distribution.

3. Models for the marginal cumulative hazards for U_i and S_i: Walker et al. (2004) assumed that marginal cumulative hazards for U_i and S_i are Weibull or Gompertz.

Walker et al. (2004)'s joint parametric model aims to remove the informative censoring problem by modeling the relation between U_i and S_i. Although the problem of informative censoring can be resolved through modeling the joint parametric model, the independent assumption between $U_i|S_i$ and C_i is problematic and difficult to check in practice since U_i is unobserved. On the other hand, the assumption of independence between T_i and C_i is standard and intuitive.

1.4.4 Switching Effects

The failure to account for the relationship between switch status and prognosis might introduce bias. Shao, Chang and Chow (2005) introduced a switching effect $\omega(s)$ in the AFT model. The inclusion of switching status accounts for the effect from potential factors related to switching status, such as disease prognosis, or other factors which makes a patient more likely to switch the regime. For the case where patients may switch from either groups, Shao et al. (2005) assumed the following model conditional on S_i,

$$U_i = S_i + e^{\psi(1-2R_i)}\omega_{R_i,\eta}(S_i)(T_i - S_i),$$

where $\omega_{R_i,\eta}(S_i)$ models the switching effect. Shao et al. (2005) suggested that function ω should be close to 1 when s is approaching 0, that is, switching effect is negligible at an early stage. This may not be appropriate since patients switch early may be sicker patients.

Assume that the survival time T_i for the control arm follows distribution $F_\theta(t)$, where θ is an unknown parameter vector. The survival time distribution can be written as $P(T_i \leq t) = F_\theta(e^{\psi R_i}t),\ t > 0$.

The distributions of the survival times conditional on S_i for patients who switched treatments are

$$
\begin{aligned}
P(T_i \le t) &= P(U_i \le S_i + e^{\psi(1-2R_i)}\omega_{R_i,\eta}(S_i)(t-S_i)) \\
&= F_\theta(e^{\psi R_i}\left[S_i + e^{\psi(1-2R_i)}\omega_{R_i,\eta}(S_i)(t-S_i)\right]) \\
&= F_\theta(e^{\psi R_i}S_i + e^{\psi(1-R_i)}\omega_{R_i,\eta}(S_i)(t-S_i)).
\end{aligned}
$$

Define $S_i = \infty$ when patient i who never switches. Then the conditional likelihood function given S_i is

$$
\begin{aligned}
L(\theta,\psi,\eta) = &\prod_{i:S_i=\infty}\left[e^{\psi R_i}f_\theta(e^{\psi R_i}Y_i)\right]^{\delta_i}\left[1-F_\theta(e^{\psi R_i}Y_i)\right]^{1-\delta_i} \\
&\times\prod_{i:S_i<\infty}\left[e^{\psi(1-R_i)}\omega_{R_i,\eta}(S_i)f_\theta\left[e^{\psi R_i}\delta_i + e^{\psi(1-R_i)}\omega_{R_i,\eta}(S_i)(Y_i-S_i)\right]\right]^{\delta_i} \\
&\times\left[1-F_\theta\left[e^{\psi R_i}S_i + e^{\psi(1-R_i)}\omega_{R_i,\eta}(S_i)(Y_i-S_i)\right]\right]^{1-\delta_i}.
\end{aligned}
$$

The use of conditional likelihood is problematic since it may lead to biased estimates (Zeng, Chen, Chen, Ibrahim and Amgen Research Group 2012).

1.5 Semi-Parametric Semicompeting Risks Transition Approach

Zeng et al. (2012) proposed to model the observed survival times using a semi-parametric hazard model. They introduced a mixed semicompeting risks

transition model to take prognosis factor into account for the disease-progression and non-disease-progression subpopulations. Time-to-death is the only event time of interest for non-disease-progression subpopulation, while both time to disease progression and event of death are considered for disease-progression subpopulation. This method of modeling event times can be viewed as an illness-death with four states: alive with/without progression, and dead with/without progression. Zeng et al. (2012) assumed Cox proportional hazard models for the time to the intermediate event and time from the intermediate event to death. One of the potential problems of the AFT model in Section 1.4 is the failure to account for the relationship between switching status and disease prognosis, which also might have impact on treatment effect. Zeng et al. (2012)'s approach not only allows the treatment effect at switching to be completely different from the baseline treatment effect, but also is able to handle both baseline covariates and prognostic covariates at switching.

1.6 Limitations of Existing Methods and Proposed Methods

Robins and Tsiatis (1991) and Branson and Whitehead (2002) proposed the accelerated failure time model (AFT) to relate the counterfactual time U_i and

observed time T_i. The relationship between switching time and survival time has not been considered. The dependence between T_i and S_i needs to be considered because a patient is more likely to switch when his/her prognosis is more severe. The conditional likelihood approach used in Shao et al. (2005) may lead to biased estimates.

Walker et al. (2004) assumed a joint parametric model for (U_i, S_i) to account for the dependence between U_i and S_i with the aim of removing the informative censoring in U_i. Although the problem of informative censoring can be resolved through modeling joint parametric model, the independence assumption between $U_i|S_i$ and C_i is problematic and difficult to check. Furthermore, the maximum likelihood does not yield randomization-respecting estimators because the original randomization balance has not been used in the estimation procedure. That is, the estimated values of U_i are not guaranteed to be balanced between randomized groups. Walker et al. (2004) obtained a randomization-respecting estimator through an adjustment procedure to ensure the randomization balance.

Zeng et al. (2012) described the different behaviors of the two subpopulations by including prognosis factors. The assumption of the treatment effect from progression depends only on the event of progression may be inappropraite. In Chapter 2, we will look at Zeng et al. (2012)'s model more closely and propose a

semi-competing risk model with a frailty term that models the relation between time to progression and time from progression to death. In Chapter 3, an AFT model will be proposed as an extension of Zeng et al. (2012)'s model. In Chapter 4, we further develop the AFT model with a frailty term.

Chapter 2

Semi-Competing Risk Models

2.1 Introduction

Zeng et al. (2012) proposed to model the observed survival times using a semi-parametric hazard model. They introduced a mixed semicompeting risks transition model to take prognosis factor into account for the disease-progression and non-disease-progression subpopulations, where disease-progression population will eventually develop disease progression before death while non-disease-progression population will never experience disease progression. Therefore, time-to-death is the only event time of interest for non-disease-progression subpopulation, and both time to disease progression and time from progression to the event of death are considered for disease-progression subpopulation.

2.2 The Semi-Computing Risk Model

2.2.1 Notations

In the remainder of this thesis, we will use the following notations:

- Indicator for progression status $P = \begin{cases} 0, & \text{no progression before death,} \\ 1, & \text{progression before death.} \end{cases}$
- Indicator for switching status $V = \begin{cases} 0, & \text{no switch,} \\ 1, & \text{switch.} \end{cases}$
- Indicator for randomized regime $R = \begin{cases} 0, & \text{control,} \\ 1, & \text{treatment.} \end{cases}$
- Indicator for censoring $\delta = \begin{cases} 0, & \text{censored,} \\ 1, & \text{uncensored.} \end{cases}$
- T_D denotes time-to-death.
- T_P denotes time to disease progression.
- T_V denote time to switch event.
- T_G stands for time from disease progression to death.
- $\mathbf{X}$ represents baseline covariates.
- $\mathbf{Z}$ reflects covariates collected at progression, as prognostic factors for the switching decisions.

- Time-dependent indicator for the switch event $U(t) = \begin{cases} 1, & \text{if } t \geq T_V, \\ 0, & \text{if } t < T_V. \end{cases}$

- $h_D(t|R, \mathbf{X}, P = 0)$, $h_P(t|R, \mathbf{X}, P = 1)$ and $h_G(t|R, \mathbf{Z}, V, P = 1, T_P)$ are the conditional hazard functions of T_D, T_P, and T_G, respectively.

- $h_{D_U}(t|U(t)), \mathbf{X}$ denotes the hazard function of T_D given time-dependent indicator for switch event $U(t)$.

- $h_0(t)$, $h_1(t)$ and $h_2(t)$ are unknown baseline hazard functions for T_D, T_P and T_G.

- $S_D(t|R, \mathbf{X}, P = 0), S_P(t|R, \mathbf{X}, P = 1)$ and $S_G(t|R, \mathbf{Z}, V, P = 1, T_P)$ are conditional survival functions of T_D, T_P, and T_G, respectively.

- $S_0(t)$, $S_1(t)$ and $S_2(t)$ are unknown baseline survival functions.

- C denotes censoring time.

- Y denotes observed failure time.

- W denotes the observed disease progression time.

2.2.2 The Model

Zeng et al. (2012)'s model is given by the following equations:

$$\text{logit}\{Pr(P=1|R,\mathbf{X})\} = \alpha_0 + \alpha_1 R + \alpha_2 \mathbf{X}, \tag{2.1}$$

$$h_D(t|R,\mathbf{X},P=0) = h_0(t)\exp\left[\beta_{00}R + \beta_{01}\mathbf{X}\right], \tag{2.2}$$

$$h_P(t|R,\mathbf{X},P=1) = h_1(t)\exp\left[\beta_{10}R + \beta_{11}\mathbf{X}\right], \tag{2.3}$$

$$h_G(t|R,\mathbf{Z},V,P=1,T_P) = h_2(t)\exp\left[\beta_{20}R + \beta_{21}V(1-R) + \beta_2^T(\mathbf{Z}^T,T_P)^T\right], \tag{2.4}$$

where $h_0(t)$, $h_1(t)$ and $h_2(t)$ are unknown baseline hazard functions, and α's and β's are regression coefficients.

Equation (2.1) models the probability of progression given baseline covariates $\mathbf{X}$ and randomized treatment R. Equation (2.2) models hazard function for non-disease-progression group given $\mathbf{X}$ and R. For disease-progression group, equation (2.3) models the hazard function of time to progression T_P given $\mathbf{X}$ and R, and equation (2.4) models the hazard function of time T_G from progression to death given V, R, $\mathbf{Z}$ and T_P. Notice that equations (2.3) and (2.4) apply to the disease-progression group only.

The ultimate aim of Zeng et al. (2012)'s model is to compare the survival function of death time to the setting when no subjects have switching. Zeng et al. (2012) adopted a counterfactual outcome framework by defining $T_D^*(a)$ for $a = 0$ as control or 1 as treatment to be a potential survival time when a subject receives treatment regime and never changes treatment status and letting $S_a(t) = P(T_D^*(a) > t)$. The goal is to compare $S_1(t)$ and $S_0(t)$. As in Robins and

Tsiatis (1991)'s causal model framework, Zeng et al. (2012) made the following assumptions.

Assumption 1. Treatment R is completely randomized and $T_D^*(a) = T_D(a)$ if a subject never changes treatment regime.

Assumption 2. Given $(R = 0, \mathbf{Z}, T_P = s)$, that is, a subject in the control arm has disease progression at time s and covariates $\mathbf{Z}$, or $(R = 1, \mathbf{Z}, T_P = s)$, V is independent of the potential outcomes $\{T_D^*(0), T_D^*(1)\}$. Denote $f_{\mathbf{X}}(\mathbf{x})$ and $f_{\mathbf{Z}}(\mathbf{z})$ as the density functions for $\mathbf{X}$ and $\mathbf{Z}$, the potential survival function of treatment a, $P(T_D^*(a) > t)$, can be obtained through integrating the conditional survival function for scenarios of $P = 0$ and $P = 1$.

In other words, $S_a(t)$ can be expressed in terms of the parameters in model (2.1), (2.2), (2.3), and (2.4), and the distribution of $\mathbf{X}$ and $\mathbf{Z}$ given $(\mathbf{X}, P = 1, R)$. All the model parameters, including α's, β's and baseline hazard functions, are estimated via nonparametric maximum likelihood. Therefore, $S_a(t)$ is estimable so that the casual effect of treatment can be estimated.

Assumption 3. The censoring time is independent of T_D, T_P and T_G given the observed covariates.

Assumption 4. For progression subjects, T_P is independent of $\mathbf{Z}$ given R and $\mathbf{X}$.

2.2.3 Limitations

Zeng et al. (2012) summarized the different behaviors of the two subpopulations through semi-parametric hazard model by including prognosis as a binary factor, depending on whether the disease of patient will progress before death time or not. Zeng et al. (2012)'s model, however, is in the form of proportional hazards with the assumption that the hazards of two individuals is related by a proportionality constant independent of t. Such an assumption might be too restrictive for some applications. Furthermore, in Zeng et al. (2012)'s model, time to progression is considered as a covariate in the Cox PH model for the time from progression event to death. An alternative approach to model the relation between time to progression and time from progression to death is to include a frailty term. We propose an accelerated failure time (AFT) model with a frailty term to account for correlation between time to progress and time from progress to death among subjects.

Chapter 3

Accelerated Failure Time Model

3.1 The Accelerated Failure Time Model

Zeng et al. (2012)'s model is in the form of proportional hazards with the assumption that the hazards of two individuals is related by a proportionality constant independent of t. Such an assumption may be restrictive for some applications. Therefore, we propose the accelerated failure time (AFT) model as alternative approach.

An accelerated failure time model assumes that the effect of given covariates, $\mathbf{Z}$, is to speed up the time by a factor $\exp\left[\theta^T\mathbf{Z}\right]$, that is, the survival function of a subject with covariates $\mathbf{Z}$ at time t is the same as the survival function of a subject with a baseline survival at a time $\exp(\theta^T\mathbf{Z})t$:

$$S(t|\mathbf{Z}) = S_0(\exp\left[\theta^T\mathbf{Z}\right] t)$$

where $S_0(t)$ is the baseline survival function with $\mathbf{Z} = 0$. In terms of the hazard function, above survival function can also be written as

$$h(t|\mathbf{Z}) = \exp\left[\theta^T \mathbf{Z}\right] h_0(\exp\left[\theta^T \mathbf{Z}\right] t).$$

We will consider the following model

$$\begin{aligned}
\text{logit}\{Pr(P = 1|R, \mathbf{X})\} &= \alpha_0 + \alpha_1 R + \alpha_2 \mathbf{X}, && (3.1)\\
S_D(t|R, \mathbf{X}, P = 0) &= S_0(\exp\left[\beta_{00} R + \beta_{01} \mathbf{X}\right] t), && (3.2)\\
S_P(t|R, \mathbf{X}, P = 1) &= S_1(\exp\left[\beta_{10} R + \beta_{11} \mathbf{X}\right] t), && (3.3)\\
S_G(t|R, \mathbf{Z}, V, P = 1, T_P) &= S_2(\exp[\beta_{20} R + \beta_{21} V(1 - R) \\
&+ \beta_2^T (\mathbf{Z}^T, T_P)^T] t), && (3.4)
\end{aligned}$$

Compare to Robins and Tsiatis (1991)'s model, the switching status $\triangle$ represents progression status P, while the switching time S represents progression time T_P, while counterfactual time could be rewritten as $U = T_P + \exp(\psi) T_G$, and time from switching to death T_G could be restated as $T_G = \exp(-\psi)(U - T_P)$. Moreover, Robins and Tsiatis (1991)'s model assumes all progressive patients switch, and there are patients with switch in the treatment arm. Specifically, Robins and Tsiatis (1991) assumes that

$$
\begin{aligned}
P(T_G > t|T_P = s) &= P(e^{-\psi}(U - s) > t|T_P = s) \\
&= P(U > e^{-\psi}t + s|T_P = s) \\
&= S_0(e^{-\psi}t + s|T_P = s). \qquad (3.5)
\end{aligned}
$$

Respectively, for the situation of no progression in treatment arm and all progressive patients switch ($R = 0$, $P = 1$ and $V = 1$), the survival distribution of time from progression to death, T_G is

$$
S_G(t|R, \mathbf{Z}, V = 1, P = 1, T_P, R = 0) = S_2(\exp[\beta_{21} + \beta_2^T(\mathbf{Z}^T, T_P)^T]t). \qquad (3.6)
$$

Therefore, Robins and Tsiatis (1991)'s model is a special case of our AFT model. In other words, when two baseline survival functions is equivalent in form of $S_0(e^{\psi}t + s|T_P = s) = S_2(\exp[\beta_{21} + \beta_2^T(\mathbf{Z}^T, T_P)])$, Robins and Tsiatis (1991)'s model stands for proposed AFT model with all progressive patients switch and no switch occurs in treatment arm.

T_D describes the survival function of time to death in equation (3.2) as progression is unobserved ($P = 0$). β_{00} stands for the accelerated factor of the treatment for patients without progression.

T_P describes the survival function of time to progression in equation (3.3) as progression is present, where β_{10} stands for the accelerated factor of the treatment on time to progression, while T_G describes the survival time from to progression to death in equation (3.4), where β_{20} stands for the accelerated factor of the treatment on time to progression to death.

When randomization regime is treatment and subject stays on treatment, β_{20} stands for the accelerated factor of the treatment, while as randomization regime is control and subject switches to treatment arm, β_{21} stands for the accelerated factor of the treatment.

In terms of the hazard function, above model can be rewritten as

$$\begin{aligned} h_D(t|R,\mathbf{X},P=0) &= h_0(\exp\left[\beta_{00}R+\beta_{01}\mathbf{X}\right]t)\exp\left[\beta_{00}R+\beta_{01}\mathbf{X}\right], \quad (3.7)\\ h_P(t|R,\mathbf{X},P=1) &= h_1(\exp\left[\beta_{10}R+\beta_{11}\mathbf{X}\right]t)\exp\left[\beta_{10}R+\beta_{11}\mathbf{X}\right], \quad (3.8)\\ h_G(t|R,\mathbf{Z},V,P=1,T_P) &= h_2(\exp\left[\beta_{20}R+\beta_{21}V(1-R)+\beta_2^T(\mathbf{Z}^T,T_P)^T\right]t)\\ &\times \exp\left[\beta_{20}R+\beta_{21}V(1-R)+\beta_2^T(\mathbf{Z}^T,T_P)^T\right]. \quad (3.9) \end{aligned}$$

We shall note, the proposed AFT model summarizes the relation between time to progression, T_P, and time from progression to death, T_G, through survival function conditional on time to progression, $S_G(t|T_P)$. An alternative approach to account for the relation between T_P and T_G will be introduced in the next chapter.

3.2 Estimation

The observed data can be divided into four groups of observations:

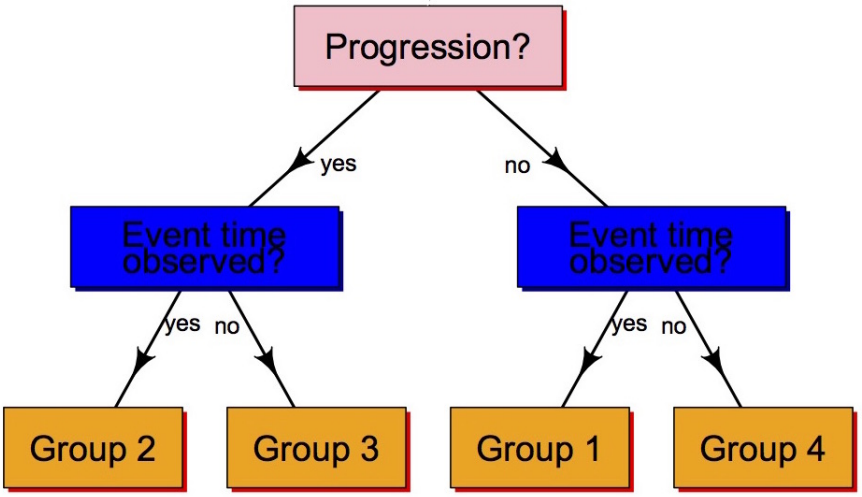

Figure 3.1: Grouping criteria by progression and censoring status shown in flowchart.

Group 1. Subjects are observed to fail at time Y without disease progression event occurring. These subjects have $P = 0, T_D = Y$ and $\delta = 1$.The contribution to the likelihood function is

$$\begin{aligned} & h_D(Y|\mathbf{X}, R) S_D(Y|\mathbf{X}, R) P(P = 0|R, \mathbf{X}) f_{\mathbf{X}}(\mathbf{x}|R) P(R) \\ = \; & h_0(\exp\left[\beta_{00} R + \beta_{01} \mathbf{X}\right] Y) \exp\left[\beta_{00} R + \beta_{01} \mathbf{X}\right] S_0(\exp\left[\beta_{00} R + \beta_{01} \mathbf{X}\right] Y) \\ & \times P(P = 0|R, \mathbf{X}) f_{\mathbf{X}}(\mathbf{x}|R) P(R), \end{aligned} \tag{3.10}$$

where $f_{\mathbf{X}}(\mathbf{x}|R=a) = \sum_{j=1}^{n} I(X_j = x, R_j = a)/\sum_{j=1}^{n} I(R_j = a)$. Because of the randomization, $f_{\mathbf{X}}(\mathbf{x}|R=a) = f_{\mathbf{X}}(\mathbf{x})$, and therefore, we can replace $\hat{f}_{\mathbf{X}}(\mathbf{x}|R=a)$ with the empirical distribution of X.

Group 2. Subjects are observed with disease progression at time W and fail at time Y. These subjects have $P=1$, $T_P = W$, $T_G = Y - W$, and $\delta = 1$. The contribution to the likelihood function is

$$\begin{aligned} & h_P(W|R,\mathbf{X},P=1)h_G(Y-W|R,V,P=1)S_P(W|R,\mathbf{X},P=1) \\ & \times S_G(Y-W|R,V,P=1)P(P=1|R,\mathbf{X})f_{\mathbf{X}}(\mathbf{x}|R)P(R) \\ = \; & h_P(W|R,\mathbf{X},P=1)h_G(Y-W|R,V,P=1)S_1(\exp\left[\beta_{10}R+\beta_{11}\mathbf{X}\right]Y) \\ & \times S_2(\exp[\beta_{20}R+\beta_{21}V(1-R)+\beta_2^T(Z^T,T_P)^T]P(P=1|R,\mathbf{X})f_{\mathbf{X}}(\mathbf{x}|R)P(R). \end{aligned} \tag{3.11}$$

Group 3. Subjects are observed with disease progression at time W and censored at time C. These subjects have $P=1$, $T_P = W$, $T_G > Y - W$ and $\delta = 0$. The contribution to the likelihood function is

$$\begin{aligned} & h_P(W|R,\mathbf{X},P=1)S_P(W|R,\mathbf{X},P=1) \\ & \times S_G(Y-W|R,V,P=1) \times P(P=1|R,\mathbf{X})f_{\mathbf{X}}(\mathbf{x}|R)P(R). \end{aligned} \tag{3.12}$$

Group 4. Subjects are only observed to be censored at time Y without disease progression event occurring. These subjects have $P = 0$, with $T_D > Y$; or, they may have $P = 1$, with $T_P > Y$. The contribution to the likelihood function is

$$
\begin{aligned}
&[S_D(Y|\mathbf{X}, R)P(P = 0|R, \mathbf{X}) + S_P(W|R, \mathbf{X}, P = 1)P(P = 1|R, \mathbf{X})]f_{\mathbf{X}}(\mathbf{x}|R)P(R) \\
= \; &[S_0(\exp[\beta_{00}R + \beta_{01}\mathbf{X}]\, Y)P(P = 0|R, \mathbf{X}) \\
&+S_1(\exp[\beta_{10}R + \beta_{11}\mathbf{X}]\, Y)P(P = 1|R, \mathbf{X})] \times f_{\mathbf{X}}(\mathbf{x}|R)P(R) \qquad (3.13)
\end{aligned}
$$

Let C_D and C_G denote the censoring time for T_D and T_G, respectively. We could write observed failure time Y and observed progression time W as

$$
\begin{aligned}
Y &= \{T_D(1 - P) + (T_P + T_G)P\}\delta + \{(C_D(1 - P) + (T_P + C_G)P\}(1 - \delta), \\
W &= \{T_D(1 - P) + T_P P\}\delta + \{C_D(1 - P) + T_P P\}(1 - \delta).
\end{aligned}
$$

The AFT model, with and without random effects, can be fitted using the SAS procedure **PROC NLMIXED**. For constructed likelihood function, the **NLMIXED** procedure finds MLEs for a range of non-linear, mixed effects models by maximizing an approximation to the full data likelihood function, integrated over the random effects. The specific distribution of random effect is defined by **RANDOM** statement in **NLMIXED**. Currently, the only distribution available for the random effects is normal with specified mean and variance. Since the proposed AFT model in this chapter does not contain random effect, the integration step is skipped here. The AFT model with random effect, or frailty, will be

introduced in Chapter 4. For detailed information about **PROC NLMIXED** procedure, please refer to Appendix A.1.

3.3 Simulation Study

With the aim to examine the performance of the proposed estimation procedure, we conducted a simulation study by generating data from (3.1), (3.2), (3.3) and (3.4). Specifically, for the first half of subjects are assigned to treatment ($R = 1$), while the rest are assigned to control ($R = 0$). We assume two baseline covariates $X_1 \overset{\text{i.i.d.}}{\sim} \text{uniform}[-1, 1]$ and $X_2 \overset{\text{i.i.d.}}{\sim} \text{Ber}(0.6)$. We consider two sample sizes: $n = 400$ and $n = 1000$.

Setting $\alpha_0 = 1.6$, $\alpha_1 = -1.8$, $\alpha_2 = 1$ and $\alpha_3 = 0.1$, the progressive status, P, is determined by (3.4). That is, P follows Bernoulli with success rate $\frac{1}{\{1+\exp[-(1.6-1.8R+X_1+0.1X_2)]\}}$. Assume that all survival times follow exponential distribution. For subjects in non-progression group with $P = 0$, the death time T_D is generated using (3.2) with $S_0(t) = \exp(-t)$, $\beta_{00} = -1$, $\beta_{01} = 1$ and $\beta_{02} = 0.2$. For subjects in progression group with $P = 1$, the time to disease progression T_P is generated using (3.3) with $S_1(t) = \exp(-t)$, $\beta_{10} = -0.5$, $\beta_{11} = 1$ and $\beta_{12} = 0$. Lastly, in order to generate the time from disease progression to death for subjects with $P = 1$, the assignment of switching indicator, V, in the untreated subjects is assumed to follow Bernoulli with success rate $\frac{1}{\{1+\exp[-(-0.1+0.3T_P+0.2X_1+0.5Z)]\}}$,

yielding a switching rate of 43.37% and 43.295% for sample size 400 and 1000 in the control arm. The time from disease progression to death, T_G, is generated using (3.4) with $S_2(t) = \exp[-t]$, $\beta_{20} = -0.3$, $\beta_{21} = -0.5$, $\beta_{22} = 0.6$, $\beta_{23} = -0.5$, $\beta_{24} = 0.5$ and $\beta_{25} = -0.4$. We set the study duration to be $\tau = 3$, and the censoring scheme is generated from a uniform distribution on $(1, 7)$; the censoring scheme is assumed to take the minimum of two.

The summary of the simulation study for 100 datasets is presented in Table 3.1, 3.2, 3.3, and 3.4. For $n = 400$, the average proportions for groups 1 to 4 are 23.92%, 45.60%, 9.58% and 20.91%, while for $n = 1000$, the average proportions for groups 1 to 4 are 23.39%, 45.80%, 9.66% and 21.16%.

The following abbreviation are used in all tables: APE represents average of the parameter estimates, SD represents sample standard deviation of the estimates, ASD represents average of the standard deviation estimates, MSE represents mean squared error, and CP represents coverage percentage of the 95% confidence interval.

From 100 simulated dataset, the obtained MSE for average of estimates are all below 0.1 and CP are close to 95% for both sample size 400 and 1000. We conclude that the proposed method performed well in estimating the true parameters.

Table 3.1: Simulation results when $n = 400$

	True	APE	Bias	SD	ASD	MSE	CP
Susceptibility Model							
α_0	1.6	1.66	0.06	0.30	0.27	0.09	94.00
α_1	-1.8	-1.88	0.08	0.31	0.33	0.10	95.00
α_2	1.0	1.05	0.05	0.28	0.31	0.08	91.00
α_3	0.1	0.06	0.04	0.29	0.29	0.09	96.00
Survival model of no-progression population							
β_{00}	-1.0	-0.97	0.03	0.21	0.24	0.09	92.00
β_{01}	1.0	0.99	0.01	0.23	0.24	0.05	95.00
β_{02}	0.2	0.23	0.03	0.23	0.26	0.05	91.00

Table 3.2: Simulation results when $n = 400$

Disease progression model of progression population

	True	APE	Bias	SD	ASD	MSE	CP
β_{10}	-0.5	-0.49	0.01	0.19	0.19	0.04	94.00
β_{11}	1.0	1.00	0.00	0.15	0.16	0.02	96.00
β_{12}	0.0	0.01	0.01	0.11	0.11	0.01	91.00

Gap time model of progression population

β_{20}	-0.3	-0.27	0.03	0.18	0.19	0.03	94.00
β_{21}	-0.5	-0.50	0.00	0.17	0.20	0.03	92.00
β_{22}	0.6	0.61	0.01	0.12	0.11	0.01	95.00
β_{23}	-0.5	-0.50	0.00	0.16	0.16	0.03	91.00
β_{24}	0.5	0.51	0.01	0.16	0.15	0.03	95.00
β_{25}	-0.4	-0.43	0.03	0.22	0.24	0.05	94.00

Table 3.3: Simulation results when $n = 1000$

	True	APE	Bias	SD	ASD	MSE	CP
Susceptibility Model							
α_0	1.6	1.64	0.04	0.30	0.27	0.09	94.00
α_1	-1.8	-1.81	0.01	0.31	0.31	0.10	95.00
α_2	1.0	1.06	0.06	0.30	0.28	0.09	91.00
α_3	0.1	0.06	0.04	0.34	0.29	0.12	96.00
Survival model of non-progression population							
β_{00}	-1.0	-0.98	0.02	0.22	0.21	0.05	92.00
β_{01}	1.0	1.06	0.06	0.25	0.25	0.07	95.00
β_{02}	0.2	0.21	0.01	0.26	0.23	0.07	91.00

Table 3.4: Simulation results when $n = 1000$

Disease progression model of progression population

	True	APE	Bias	SD	ASD	MSE	CP%
β_{10}	-0.5	-0.52	0.02	0.18	0.19	0.03	94.00
β_{11}	1.0	1.00	0.00	0.15	0.15	0.01	96.00
β_{12}	0.0	0.04	0.04	0.13	0.11	0.02	91.00

Gap time model of progression population

β_{20}	-0.3	-0.30	0.00	0.15	0.19	0.02	94.00
β_{21}	-0.5	-0.51	0.01	0.18	0.17	0.01	92.00
β_{22}	0.6	0.57	0.03	0.10	0.09	0.01	95.00
β_{23}	-0.5	-0.52	0.02	0.13	0.14	0.02	91.00
β_{24}	0.5	0.52	0.02	0.14	0.14	0.02	95.00
β_{25}	-0.4	-0.37	0.03	0.23	0.22	0.05	94.00

Chapter 4

Accelerated Failure Time Model with Frailty

4.1 The Accelerated Failure Time Model with Frailty

In Chapter 3, we proposed AFT model as an extension of Zeng et al. (2012)'s model. Time to progression is considered as a covariate. An alternative approach to account for relation between time to progression and time from progression to death is to introduce a frailty.

4.2 Bivariate Frailty Model

Shared frailty is often used to model association between event times. When two observed survival times depend on the same unobserved variable, called a frailty, this common dependence induces an association between the observed

times.

In (3.4), the relation between time to progression, T_P, and time from progression to death, T_G is modeled through survival function of T_G condition on T_P. Frailty model, as an alternative approach to address this relation is considered here. We propose a frailty model to account for the relation between time to progression, T_P, and time from progression to death, T_G. Given the common frailty ω, we assume that T_P and T_G are independent. Consider a general survival time variable T in log-linear form with a frailty term, $-\omega$, as additive random factor, is expressed as

$$\log T = \mu + \theta^T \mathbf{X} + (-\omega) + \sigma\epsilon, \tag{4.1}$$

where μ and σ are unknown location and scale parameters, while ϵ has some specific probability distribution.

Let $\psi(\mathbf{X}, \omega|\theta) = \exp\left[\theta^T \mathbf{X} + (-\omega)\right]$. The survival function of T,

$$\begin{aligned} S(t|\mathbf{X}, \omega) &= P(T \geq t|\mathbf{X}, \omega) \\ &= P(\log T \geq \log t|\mathbf{X}, \omega) \\ &= P(\mu + \sigma\epsilon \geq \ln t - \theta^T \mathbf{X} + \omega) \\ &= P(\exp\left[\mu + \sigma\epsilon\right] \geq t/\psi(\mathbf{X}, \omega)). \end{aligned} \tag{4.2}$$

The baseline survival function, $S_0(t)$, the survival function of $\mathbf{X} = \mathbf{0}$, is

$$S_0(t) \quad = \quad P(\exp\left[\mu + \sigma\epsilon\right] \geq t). \tag{4.3}$$

Conditional on ω, the AFT model is assumed to hold,

$$S(t) \quad = \quad S_0(\exp\left[\beta^T\mathbf{X} + \omega\right] t). \tag{4.4}$$

where $\theta = -\beta$. The corresponding hazard function is (Collett 2003)

$$h(t) \quad = \quad \exp\left[\beta^T\mathbf{X} + \omega\right] \cdot h_0(\exp\left[\beta^T\mathbf{X} + \omega\right] t). \tag{4.5}$$

Now, considering T_P and T_G in log-linear form with a frailty term ω, and T_D in the same form without a frailty term as

$$\log T_D \quad = \quad \mu_D + \theta_{00}R + \theta_{01}\mathbf{X} + \sigma_D\epsilon_D, \tag{4.6}$$

$$\log T_P \quad = \quad \mu_P + \theta_{10}R + \theta_{11}\mathbf{X} + (-\omega) + \sigma_P\epsilon_P, \tag{4.7}$$

$$\log T_G \quad = \quad \mu_G + \theta_{20}R + \theta_{21}V(1-R) + \theta_2^T(\mathbf{Z}^T, T_P)^T + \theta_3(-\omega) + \sigma_G\epsilon_G, \tag{4.8}$$

where $\mu_D, \mu_P, \mu_G, \sigma_D, \sigma_U$ and σ_G are unknown location and scale parameters, while ϵ_D, ϵ_P and ϵ_G have some specific probability distributions. Given frailty ω, AFT model of T_D, T_P and T_G are assumed to be

$$
\begin{aligned}
S_D(t|R,\mathbf{X},P=0) &= S_0(\exp\left[\beta_{00}R+\beta_{01}\mathbf{X}\right]t), && (4.9)\\
S_{P,G}(t_p,t_g|R,\mathbf{X},\mathbf{Z},P=1,\omega) &= S_P(t_p|R,\mathbf{X},U=1,\omega)\\
&\quad\times\ S_G(t_g|R,\mathbf{Z},V,P=1,\omega), && (4.10)\\
&= S_1(\exp\left[\beta_{10}R+\beta_{11}\mathbf{X}+\omega\right]t_p)\\
&\quad\times\ S_2(\exp[\beta_{20}R+\beta_{21}V(1-R)\\
&\quad+\ \beta_2^T(\mathbf{Z}^T,T_P)^T+\beta_3\omega]t_g), && (4.11)\\
h_D(t|R,\mathbf{X},P=0) &= \exp\left[\beta_{00}R+\beta_{01}\mathbf{X}\right]h_0(\exp.\left[\beta_{00}R+\beta_{01}\mathbf{X}\right]t), && (4.12)\\
h_P(t_p|R,\mathbf{X},P=1,\omega) &= \exp\left[\beta_{10}R+\beta_{11}\mathbf{X}+\omega\right]\\
&\quad\times\ h_1(\exp\left[\beta_{10}R+\beta_{11}\mathbf{X}+\omega\right]t_p), && (4.13)\\
h_G(t_g|R,\mathbf{Z},P=1,\omega) &= \exp\left[\beta_{20}R+\beta_{21}V(1-R)+\beta_2^T(\mathbf{Z}^T,T_P)^T+\omega\right]\\
&\quad\times\ h_2(\exp[\beta_{20}R+\beta_{21}V(1-R)\\
&\quad+\ \beta_2^T(\mathbf{Z}^T,T_P)^T+\beta_3\omega]]t_g). && (4.14)
\end{aligned}
$$

where $S_{P,G}(t_p,t_g|R,\mathbf{X},\mathbf{Z},P=1,\omega)$ denotes joint survival function of time to progression and time to progression to death given a frailty term and relevant covariates when progression is present. Note $\theta_{00}=-\beta_{00}$, $\theta_{01}=-\beta_{01}$, $\theta_{10}=-\beta_{10}$, $\theta_{20}=-\beta_{21}$, and $\theta_3=-\beta_3$. Note T_P and T_G are assumed to be independent given ω. Also, T_P in (4.11) are considered to be prognosis factor.

A logit model is used to determine the progression status,

$$\text{logit}\{Pr(P=1|R,\mathbf{X})\} \quad = \quad \alpha_0 + \alpha_1 R + \alpha_2 \mathbf{X}, \tag{4.15}$$

We assume that ω are i.i.d. $\mathcal{N}(0,\sigma^2)$.

When there is no progression ($P=0$), equation (4.9) describes the survival function of time to death, T_D, where β_{00} stands for the accelerated factor of the treatment.

When there is a progression ($P=1$), the left hand side of equation (4.10) describes the joint survival function of time to progression, T_P, and time from progression to death, T_G, given frailty term, ω, the right hand side of equation describe the product of survival function of T_U and T_G. The frailty ω is a latent variable reflecting health condition for each subject.

Furthermore, the first part and second part of equation (4.11) describes the survival function of time to progression, T_P, and the survival function of time from progression to death, T_G. In the first part, β_{10} stands for the accelerated factor of the treatment before progression. The second part addresses the time after progression event: when randomized regime is treatment and subject stays on treatment, β_{20} stands for the accelerated factor of the treatment after progression. β_{21} stands for the accelerated factor of the treatment when patients switch from

control to treatment. β_3 models the different post-progression impact on survival time among subjects, as the scaling parameter for frailty term in function for T_G.

4.3 Estimation

The observed data can be divided into four groups of observations:

Group 1. Subjects are observed to fail at time Y without disease progression event occurring. These subjects have $U = 0, T_D = Y$ and $\delta = 1$.The contribution to the likelihood function is

$$\begin{aligned} & h_D(Y|\mathbf{X}, R) S_D(Y|\mathbf{X}, R) P(P = 0|R, \mathbf{X}) f_{\mathbf{X}}(\mathbf{x}|R) P(R) \\ = \; & h_0(\exp\left[\beta_{00} R + \beta_{01}\mathbf{X}\right] Y) \exp\left[\beta_{00} R + \beta_{01}\mathbf{X}\right] \\ & \times S_0(\exp\left[\beta_{00} R + \beta_{01}\mathbf{X}\right] Y) P(P = 0|R, \mathbf{X}) f_{\mathbf{X}}(\mathbf{x}|R) P(R). \end{aligned} \tag{4.16}$$

Group 2. Subjects are observed with disease progression at time W and fail at time Y. These subjects have $P = 1$, $T_P = W$, $T_G = Y - W$ and $\delta = 1$. The contribution to the likelihood function is

$$
\begin{aligned}
& h_P(W|R,\mathbf{X},P=1,\omega)h_G(Y-W|R,V,P=1,\omega) \\
& \times S_{P,G}(W,Y-W|R,\mathbf{X},\omega)P(P=1|R,\mathbf{X})f_{\mathbf{X}}(\mathbf{x}|R)P(R) \\
= \; & h_P(W|R,\mathbf{X},PU=1,\omega)h_G(Y-W|R,V,P=1,\omega) \\
& \times S_P(W|R,\mathbf{X},P=1,\omega)S_G(Y-W|R,V,P=1,\omega) \\
& \times P(P=1|R,\mathbf{X})f_{\mathbf{X}}(\mathbf{x}|R)P(R). \qquad (4.17)
\end{aligned}
$$

Group 3. Subjects are observed with disease progression at time W and censored at time C. These subjects have $P = 1$, $T_P = W$, $T_G > Y - W$ and $\delta = 0$. The contribution to the likelihood function is

$$
\begin{aligned}
& h_P(W|R,\mathbf{X},P=1,\omega)S_P(W|R,\mathbf{X},P=1,\omega) \\
& \times S_G(Y-W|R,V,P=1,\omega) \times P(P=1|R,\mathbf{X})f_{\mathbf{X}}(\mathbf{x}|R)P(R). \qquad (4.18)
\end{aligned}
$$

Group 4. Subjects are only observed to be censored at time Y without disease progression and without death. These subjects have $P = 0$, $T_D > Y$; or, they may have $P = 1$, $T_P > Y$. The contribution to the likelihood function is

$$
\begin{aligned}
& \{S_D(Y|\mathbf{X},R)P(P=0|R,\mathbf{X}) + S_P(W|R,\mathbf{X},P=1,\omega)P(P=1|R,\mathbf{X})\} \\
& f_{\mathbf{X}}(\mathbf{x}|R)P(R) \\
= \; & [S_0(\exp[\beta_{00}R + \beta_{01}\mathbf{X}]\,Y)P(P=0|R,\mathbf{X}) \\
& + S_1(\exp[\beta_{10}R + \beta_{11}\mathbf{X} + \omega]\,Y|\omega)P(P=1|R,\mathbf{X})]f_{\mathbf{X}}(\mathbf{x}|R)P(R). \qquad (4.19)
\end{aligned}
$$

Let C_D and C_G denote the censoring death time for T_D and T_G, respectively. We could write observed failure time Y and observed progression time W as

$$Y = \{T_D(1-P) + (T_P + T_G)U\}\delta + \{(C_D(1-P) + (T_P + C_G)P\}(1-\delta)$$

$$W = \{T_D(1-P) + T_P P\}\delta + \{C_D(1-P) + T_P P\}(1-\delta)$$

As in Chapter 3, the AFT model with frailty term (random effects) can be fitted using the SAS procedure **PROC NLMIXED**. For constructed likelihood function, the **NLMIXED** procedure finds MLEs for a range of non-linear, mixed effects models by maximizing an approximation to the full data likelihood function, integrated over the random effects. The integration method is specified to take default method, adaptive Gauss-Hermite quadrature (**GAUSS**). The specific distribution of random effect is defined by **RANDOM** statement in **NLMIXED**. Currently, the only distribution available for the random effects is normal with specified mean and variance. Appendix A.2 gives an example of **PROC NLMIXED** procedure.

4.4 Simulation Study

With the aim to examine the performance of the proposed estimation procedure, we conducted a simulation study by generating data from (4.9), (4.10), (4.11) and (4.15). Specifically, the first half of subjects are assigned to treatment

$(R = 1)$, while the rest are assigned to control $(R = 0)$. We assume two baseline covariates $X_1 \overset{\text{i.i.d.}}{\sim} \text{uniform}[-1, 1]$ and $X_2 \overset{\text{i.i.d.}}{\sim} \text{Ber}(0.6)$.

Setting $\alpha_0 = 1.6, \alpha_1 = -1.8, \alpha_2 = 1$ and $\alpha_3 = 0.1$, the progressive status, U, is determined by (4.15). That is, P follows Bernoulli with success rate $\frac{1}{\{1+\exp[-(1.6-1.8R+X_1+0.1X_2)]\}}$. Assume all survival times follow exponential distribution. For subjects in non-progression group with $P = 0$, the death time T_D is generated using (4.9) with $S_0(t) = \exp(-t)$, $\beta_{00} = -1$, $\beta_{01} = 1$ and $\beta_{02} = 0.2$. For subjects in progression group with $P = 1$, the time to disease progression T_P is generated using (4.11) with $S_1(t) = \exp(-t)$, $\beta_{10} = -0.5$, $\beta_{11} = 1$ and $\beta_{12} = 0$. Lastly, in order to generate the time from disease progression to death for subjects with $U = 1$, the assignment of switching indicator, V, in the untreated subjects is assumed to follow Bernoulli with success rate $\frac{1}{\{1+\exp[-(-0.1+0.3T_P+0.2X_1+0.5Z)]\}}$. The time from disease progression to death, T_G, is generated using (4.11) with $S_2(t) = \exp[-t]$, $\beta_{20} = -0.3$, $\beta_{21} = -0.5$, $\beta_{22} = 0.6$, $\beta_{23} = -0.5$, $\beta_{24} = 0.5$, $\beta_{25} = -0.4$ and $\beta_3 = 0.2$. We set the study duration to be $\tau = 3$, and the censoring scheme is generated from a uniform distribution on $(1, 7)$; the censoring scheme is assumed to take the minimum of two.

The frailty term, ω, are i.i.d $\mathcal{N}(0, \sigma^2)$. We consider three scenarios of variances, as $\sigma^2 = 0.25, 1, 2.25$; for each scenario, sample sizes $n = 400$ and $n = 1000$

are assumed. The switching rate in the control arm for each scenarios of variances are (42.78%, 47.688%), (40.205%,40.218%), (37.88%, 37.542%). The results from 100 replicates are given in Tables 4.1 - 4.12. For $n = 400$, the average proportions for group 1 to 4 are (23.41%, 44.70%, 9.52%, 22.38%), (23.51%, 42.86%, 8.64%, 25.00%), (23.41%, 40.96%, 7.75%, 27.89%), while for $n = 1000$, the average proportions for group 1 to 4 are (23.52%, 44.95%, 9.28%, 22.26%), (23.46%,42.99%, 8.48%, 25.08%), (23.48%, 41.18%, 7.71%, 27.74%).

Similarly, as a special case that impact from subject's health condition on time of pre-progression and post-progression is the same, β_3 is assumed to be 1 and we do not estimate β_3. The switching rate in the control arm for each scenarios of variances are (41.34%,41.02%), (39.7%, 40.682%), (37.79%, 37.976%). The results from 100 replicates are given in Tables 4.13 - 4.24. For $n = 400$, the average proportions for group 1 to 4 of each case are (23.22%, 44.25%,10.05%, 22.40%), (23.27%,41.93%,9.7%, 25.10%), and (23.52%, 40.19%, 8.69%, 27.61%), while for $n = 1000$, the average proportions for group 1 to 4 of each case are (23.33%, 44.41%, 9.71%, 22.54%), (23.63%,41.78%,9.6%, 24.96%) and (23.25%,40.00%, 8.98%, 27.70%). Overall our estimation procedure performed very well.

From 100 simulated datasets, the obtained MSE for average of estimates are all below 0.5 and CP are close to 95% for both sample size 400 and 1000 when

Table 4.1: Simulation results when $\sigma^2 = 0.25$ and $n = 400$.

	True	APE	Bias	SD	ASD	MSE	CP
σ^2	0.25	0.21	0.04	0.12	0.15	0.02	93.26
Susceptibility Model							
α_0	1.6	1.62	0.02	0.24	0.27	0.06	98.96
α_1	-1.8	-1.83	0.03	0.29	0.32	0.09	97.92
α_2	1.0	1.01	0.01	0.28	0.28	0.08	94.79
α_3	0.1	0.08	0.02	0.27	0.29	0.07	94.79
Survival model of no-progression population							
β_{00}	-1.0	-0.99	0.01	0.21	0.22	0.04	94.79
β_{01}	1.0	1.05	0.05	0.24	0.24	0.06	90.63
β_{02}	0.2	0.19	0.01	0.20	0.24	0.04	98.96

$\sigma^2 = 0.25, 1, 2.25$. We conclude that the proposed estimation method performed well.

Table 4.2: Simulation results when $\sigma^2 = 0.25$ and $n = 400$.

	True	APE	Bias	SD	ASD	MSE	CP
Disease progression model of progression population							
β_{10}	-0.5	-0.47	0.03	0.22	0.22	0.05	91.67
β_{11}	1.0	0.99	0.01	0.28	0.18	0.08	95.83
β_{12}	0.0	-0.01	0.01	0.13	0.13	0.02	94.79
Gap time model of progression population							
β_{20}	-0.3	-0.31	0.01	0.20	0.20	0.04	93.75
β_{21}	-0.5	-0.51	0.01	0.17	0.18	0.03	97.92
β_{22}	0.6	0.60	0.00	0.14	0.14	0.02	93.75
β_{23}	-0.5	-0.51	0.01	0.15	0.16	0.02	96.86
β_{24}	0.5	0.52	0.02	0.14	0.15	0.02	97.92
β_{25}	-0.4	-0.40	0.00	0.23	0.24	0.05	95.83
β_3	0.2	0.27	0.07	0.68	0.71	0.47	94.00

Table 4.3: Simulation results when $\sigma^2 = 0.25$ and $n = 1000$.

	True	APE	Bias	SD	ASD	MSE	CP
σ^2	0.25	0.23	0.02	0.11	0.1	0.01	92.71
Susceptibility Model							
α_0	1.6	1.60	0.00	0.16	0.17	0.03	95.83
α_1	-1.8	-1.86	0.06	0.20	0.20	0.04	94.79
α_2	1.0	1.02	0.02	0.18	0.18	0.03	96.88
α_3	0.1	0.11	0.01	0.18	0.18	0.03	95.83
Survival model of no-progression population							
β_{00}	-1.0	-1.01	0.01	0.15	0.13	0.02	93.75
β_{01}	1.0	1.03	0.03	0.13	0.15	0.02	98.96
β_{02}	0.2	0.20	0.00	0.16	0.14	0.03	91.67

Table 4.4: Simulation results when $\sigma^2 = 0.25$ and $n = 1000$.

Disease progression model of progression population

	True	APE	Bias	SD	ASD	MSE	CP
β_{10}	-0.5	-0.44	0.06	0.15	0.15	0.03	92.71
β_{11}	1.0	0.98	0.02	0.12	0.11	0.01	93.75
β_{12}	0.0	0.00	0.00	0.09	0.08	0.01	93.75

Gap time model of progression population

β_{20}	-0.3	-0.28	0.02	0.12	0.13	0.01	94.79
β_{21}	-0.5	-0.49	0.01	0.12	0.11	0.01	92.71
β_{22}	0.6	0.59	0.01	0.01	0.01	0.00	93.75
β_{23}	-0.5	-0.53	0.03	0.11	0.10	0.01	93.75
β_{24}	0.5	0.51	0.01	0.01	0.01	0.00	94.79
β_{25}	-0.4	-0.39	0.01	0.15	0.15	0.02	96.88
β_3	0.2	0.20	0.00	0.50	0.43	0.25	89.58

Table 4.5: Simulation results when $\sigma^2 = 1$ and $n = 400$.

	True	APE	Bias	SD	ASD	MSE	CP
σ^2	1.0	0.92	0.08	0.27	0.28	0.08	97.67
Susceptibility Model							
α_0	1.6	1.65	0.05	0.31	0.27	0.10	96.51
α_1	-1.8	-1.93	0.13	0.34	0.32	0.13	95.35
α_2	1.0	1.04	0.04	0.30	0.29	0.09	94.19
α_3	0.1	0.11	0.01	0.29	0.29	0.08	95.35
Survival model of no-progression population							
β_{00}	-1.0	-1.01	0.01	0.25	0.22	0.06	91.86
β_{01}	1.0	0.99	0.01	0.27	0.25	0.07	94.19
β_{02}	0.2	0.18	0.02	0.30	0.24	0.09	86.05

Table 4.6: Simulation results when $\sigma^2 = 1$ and $n = 400$.

	True	APE	Bias	SD	ASD	MSE	CP
Disease progression model of progression population							
β_{10}	-0.5	-0.47	0.03	0.29	0.27	0.09	89.53
β_{11}	1.0	0.97	0.03	0.23	0.22	0.05	94.19
β_{12}	0.0	0.01	0.01	0.17	0.16	0.03	86.05
Gap time model of progression population							
β_{20}	-0.3	-0.29	0.01	0.19	0.21	0.04	95.35
β_{21}	-0.5	-0.51	0.01	0.16	0.19	0.03	98.84
β_{22}	0.6	0.58	0.02	0.14	0.16	0.02	96.67
β_{23}	-0.5	-0.54	0.04	0.15	0.16	0.03	95.35
β_{24}	0.5	0.51	0.01	0.16	0.16	0.03	97.67
β_{25}	-0.4	-0.34	0.06	0.23	0.25	0.06	97.67
β_3	0.2	0.18	0.02	0.27	0.24	0.07	90.70

Table 4.7: Simulation results when $\sigma^2 = 1$ and $n = 1000$.

	True	APE	Bias	SD	ASD	MSE	CP
σ^2	1.0	0.99	0.01	0.16	0.18	0.03	97.37
Susceptibility Model							
α_0	1.6	1.60	0.00	0.18	0.17	0.03	94.74
α_1	-1.8	-1.84	0.04	0.19	0.20	0.04	96.05
α_2	1.0	1.01	0.01	0.17	0.18	0.03	100.00
α_3	0.1	0.11	0.01	0.18	0.18	0.03	94.74
Survival model of no-progression population							
β_{00}	-1.0	-1.02	0.02	0.15	0.14	0.02	94.74
β_{01}	1.0	1.03	0.03	0.16	0.15	0.03	94.74
β_{02}	0.2	0.21	0.01	0.14	0.15	0.02	94.74

Table 4.8: Simulation results when $\sigma^2 = 1$ and $n = 1000$.

Disease progression model of progression population

	True	APE	Bias	SD	ASD	MSE	CP
β_{10}	-0.5	-0.47	0.03	0.17	0.18	0.03	93.42
β_{11}	1.0	0.99	0.01	0.13	0.14	0.02	93.42
β_{12}	0.0	0.00	0.00	0.09	0.10	0.01	94.74

Gap time model of progression population

β_{20}	-0.3	-0.28	0.02	0.13	0.13	0.02	93.42
β_{21}	-0.5	-0.50	0.00	0.13	0.18	0.02	93.42
β_{22}	0.6	0.57	0.03	0.10	0.11	0.01	94.74
β_{23}	-0.5	-0.50	0.00	0.09	0.10	0.03	96.05
β_{24}	0.5	0.51	0.01	0.1	0.1	0.03	93.42
β_{25}	-0.4	-0.37	0.03	0.15	0.16	0.03	94.74
β_3	0.2	0.19	0.01	0.16	0.15	0.03	90.70

Table 4.9: Simulation results when $\sigma^2 = 2.25$ and $n = 400$.

	True	APE	Bias	SD	ASD	MSE	CP
σ^2	2.25	2.15	0.10	0.48	0.46	0.24	89.41
Susceptibility Model							
α_0	1.6	1.62	0.02	0.25	0.27	0.06	98.82
α_1	-1.8	-1.85	0.05	0.35	0.33	0.13	95.30
α_2	1.0	1.07	0.07	0.28	0.29	0.08	96.47
α_3	0.1	0.09	0.01	0.27	0.30	0.07	97.65
Survival model of no-progression population							
β_{00}	-1.0	-1.00	0.00	0.27	0.23	0.07	92.94
β_{01}	1.0	1.05	0.05	0.25	0.25	0.07	98.82
β_{02}	0.2	0.15	0.05	0.27	0.24	0.08	95.30

Table 4.10: Simulation results when $\sigma^2 = 2.25$ and $n = 400$.

	True	APE	Bias	SD	ASD	MSE	CP
Disease progression model of progression population							
β_{10}	-0.5	-0.49	0.01	0.33	0.34	0.11	91.76
β_{11}	1.0	0.99	0.01	0.27	0.27	0.07	94.12
β_{12}	0.0	0.03	0.03	0.22	0.20	0.05	91.77
Gap time model of progression population							
β_{20}	-0.3	-0.29	0.01	0.21	0.21	0.04	94.12
β_{21}	-0.5	-0.51	0.01	0.17	0.20	0.03	98.82
β_{22}	0.6	0.61	0.01	0.17	0.17	0.03	92.94
β_{23}	-0.5	-0.51	0.01	0.16	0.16	0.03	94.12
β_{24}	0.5	0.51	0.01	0.16	0.16	0.03	98.82
β_{25}	-0.4	-0.36	0.04	0.26	0.26	0.07	94.12
β_3	0.2	0.19	0.01	0.13	0.14	0.02	95.29

Table 4.11: Simulation results when $\sigma^2 = 2.25$ and $n = 1000$.

	True	APE	Bias	SD	ASD	MSE	CP
σ^2	2.25	2.26	0.01	0.25	0.29	0.06	98.63
Susceptibility Model							
α_0	1.6	1.59	0.01	0.18	0.17	0.03	93.15
α_1	-1.8	-1.79	0.01	0.21	0.20	0.04	93.15
α_2	1.0	1.00	0.00	0.18	0.18	0.03	91.78
α_3	0.1	0.12	0.02	0.21	0.19	0.04	90.41
Survival model of no-progression population							
β_{00}	-1.0	-1.00	0.00	0.18	0.15	0.03	84.93
β_{01}	1.0	1.00	0.00	0.17	0.17	0.03	94.52
β_{02}	0.2	0.24	0.04	0.18	0.15	0.03	86.30

Table 4.12: Simulation results when $\sigma^2 = 2.25$ and $n = 1000$.

Disease progression model of progression population

	True	APE	Bias	SD	ASD	MSE	CP
β_{10}	-0.5	-0.52	0.02	0.24	0.22	0.06	91.78
β_{11}	1.0	1.01	0.01	0.17	0.17	0.03	91.78
β_{12}	0.0	0.02	0.02	0.10	0.13	0.01	97.26

Gap time model of progression population

β_{20}	-0.3	-0.30	0.00	0.14	0.13	0.02	94.52
β_{21}	-0.5	-0.49	0.01	0.13	0.18	0.02	93.15
β_{22}	0.6	0.59	0.01	0.11	0.10	0.01	91.78
β_{23}	-0.5	-0.51	0.01	0.11	0.10	0.01	90.41
β_{24}	0.5	0.50	0.00	0.09	0.10	0.01	98.63
β_{25}	-0.4	-0.35	0.05	0.14	0.16	0.02	98.63
β_3	0.2	0.18	0.02	0.08	0.09	0.01	95.89

Table 4.13: Simulation results when $\sigma^2 = 0.25$ and $n = 400$ with $\beta_3 = 1$.

	true	APE	bias	SD	ASD	MSE	CP
σ^2	0.25	0.24	0.01	0.11	0.12	0.01	94.00
Susceptibility Model							
α_0	1.6	1.64	0.04	0.30	0.27	0.09	93.00
α_1	-1.8	-1.85	0.05	0.37	0.32	0.14	91.00
α_2	1	0.98	0.02	0.28	0.29	0.08	97.00
α_3	0.1	0.11	0.01	0.27	0.29	0.07	99.00
Survival model of no-progression population							
β_{00}	-1	-0.95	0.05	0.25	0.22	0.07	95.00
β_{01}	1	1.01	0.01	0.27	0.24	0.07	93.00
β_{02}	0.2	0.15	0.05	0.26	0.24	0.06	95.00

Table 4.14: Simulation results when $\sigma^2 = 0.25$ and $n = 400$ with $\beta_3 = 1$.

Disease progression model of progression population

	true	APE	bias	SD	ASD	MSE	CP
β_{10}	-0.5	-0.49	0.01	0.22	0.22	0.05	91.00
β_{11}	1	1.02	0.02	0.15	0.18	0.02	99.00
β_{12}	0	0.01	0.01	0.12	0.13	0.01	95.00

Gap time model of progression population

β_{20}	-0.3	-0.28	0.02	0.24	0.22	0.06	95.00
β_{21}	-0.5	-0.49	0.01	0.18	0.20	0.03	97.00
β_{22}	0.6	0.58	0.02	0.14	0.14	0.02	97.00
β_{23}	-0.5	-0.49	0.01	0.16	0.17	0.03	97.00
β_{24}	0.5	0.50	0.00	0.16	0.16	0.03	95.00
β_{25}	-0.4	-0.41	0.01	0.27	0.25	0.07	92.00

Table 4.15: Simulation results when$\sigma^2 = 0.25$ and $n = 1000$ with $\beta_3 = 1$.

	true	APE	bias	SD	ASD	MSE	CP
σ^2	0.25	0.22	0.03	0.07	0.07	0.01	95.50
Susceptibility Model							
α_0	1.6	1.64	0.04	0.18	0.17	0.03	95.51
α_1	-1.8	-1.86	0.06	0.22	0.20	0.05	94.38
α_2	1	1.00	0.00	0.21	0.18	0.04	92.13
α_3	0.1	0.08	0.02	0.21	0.18	0.04	89.89
Survival model of no-progression population							
β_{00}	-1	-1.02	0.02	0.14	0.14	0.02	94.38
β_{01}	1	0.99	0.01	0.17	0.15	0.03	95.51
β_{02}	0.2	0.20	0.00	0.14	0.14	0.02	96.63

Table 4.16: Simulation results when $\sigma^2 = 0.25$ and $n = 1000$ with $\beta_3 = 1$.

Disease progression model of progression population

	true	APE	bias	SD	ASD	MSE	CP
β_{10}	-0.5	-0.47	0.03	0.16	0.14	0.03	94.38
β_{11}	1	0.98	0.02	0.12	0.11	0.01	95.51
β_{12}	0	0.00	0.00	0.08	0.08	0.06	96.63

Gap time model of progression population

β_{20}	-0.3	-0.26	0.04	0.15	0.13	0.02	91.01
β_{21}	-0.5	-0.50	0.00	0.12	0.12	0.01	98.88
β_{22}	0.6	0.57	0.03	0.08	0.09	0.01	95.51
β_{23}	-0.5	-0.50	0.00	0.11	0.10	0.01	93.26
β_{24}	0.5	0.50	0.00	0.10	0.10	0.01	97.75
β_{25}	-0.4	-0.36	0.04	0.17	0.16	0.03	91.01

Table 4.17: Simulation Results when $\sigma^2 = 1$ and $n = 400$ with $\beta_3 = 1$.

	True	APE	Bias	SD	ASD	MSE	CP
σ^2	1.0	0.99	0.01	0.22	0.23	0.05	97.87
Susceptibility Model							
α_0	1.6	1.64	0.04	0.25	0.27	0.06	97.87
α_1	-1.8	-1.81	0.01	0.34	0.32	0.12	91.49
α_2	1.0	1.01	0.01	0.30	0.29	0.09	96.81
α_3	0.1	0.09	0.01	0.30	0.29	0.09	93.62
Survival model of no-progression population							
β_{00}	-1.0	-0.98	0.02	0.22	0.22	0.05	93.62
β_{01}	1.0	0.98	0.02	0.24	0.25	0.06	96.81
β_{02}	0.2	0.22	0.02	0.23	0.24	0.05	91.49

Table 4.18: Simulation Results when $\sigma^2 = 1$ and $n = 400$ with $\beta_3 = 1$.

Disease progression model of progression population

	True	APE	Bias	SD	ASD	MSE	CP
β_{10}	-0.5	-0.50	0.00	0.28	0.26	0.08	94.68
β_{11}	1.0	1.04	0.04	0.22	0.22	0.05	94.68
β_{12}	0.0	0.01	0.01	0.16	0.16	0.03	91.49

Gap time model of progression population

β_{20}	-0.3	-0.34	0.04	0.27	0.27	0.07	92.55
β_{21}	-0.5	-0.52	0.02	0.24	0.24	0.06	95.74
β_{22}	0.6	0.61	0.01	0.16	0.17	0.03	96.80
β_{23}	-0.5	-0.50	0.00	0.23	0.21	0.05	93.62
β_{24}	0.5	0.50	0.00	0.17	0.20	0.03	98.94
β_{25}	-0.4	-0.38	0.02	0.30	0.31	0.10	97.87

Table 4.19: Simulation Results when $\sigma^2 = 1$ and $n = 1000$ with $\beta 3 = 1$.

	True	APE	Bias	SD	ASD	MSE	CP
σ^2	1.0	0.97	0.03	0.15	0.14	0.02	93.02
Susceptibility Model							
α_0	1.6	1.61	0.01	0.17	0.17	0.03	94.19
α_1	-1.8	-1.82	0.02	0.21	0.20	0.04	95.35
α_2	1.0	0.99	0.01	0.18	0.18	0.03	95.35
α_3	0.1	0.10	0.00	0.18	0.18	0.03	96.51
Survival model of no-progression population							
β_{00}	-1.0	-0.99	0.01	0.12	0.14	0.01	98.83
β_{01}	1.0	1.03	0.03	0.15	0.15	0.02	97.67
β_{02}	0.2	0.23	0.03	0.12	0.15	0.02	98.83

Table 4.20: Simulation Results when $\sigma^2 = 1$ and $n = 1000$ with $\beta_3 = 1$.

Disease progression model of progression population

	True	APE	Bias	SD	ASD	MSE	CP
β_{10}	-0.5	-0.53	0.03	0.19	0.17	0.04	91.86
β_{11}	1.0	1.00	0.00	0.14	0.13	0.02	97.67
β_{12}	0.0	-0.00	0.00	0.09	0.10	0.01	98.83

Gap time model of progression population

β_{20}	-0.3	-0.30	0.00	0.19	0.17	0.04	91.86
β_{21}	-0.5	-0.50	0.00	0.14	0.15	0.02	98.83
β_{22}	0.6	0.60	0.00	0.10	0.10	0.01	98.83
β_{23}	-0.5	-0.51	0.01	0.14	0.13	0.02	96.51
β_{24}	0.5	0.52	0.02	0.13	0.13	0.02	96.51
β_{25}	-0.4	-0.43	0.03	0.20	0.19	0.04	94.19

Table 4.21: Simulation Results when $\sigma^2 = 2.25$ and $n = 400$ with $\beta_3 = 1$.

	true	APE	bias	SD	ASD	MSE	CP
σ^2	2.25	2.22	0.03	0.34	0.38	0.12	97.80
Susceptibility Model							
α_0	1.6	1.65	0.05	0.29	0.27	0.09	93.41
α_1	-1.8	-1.84	0.04	0.35	0.33	0.12	92.31
α_2	1	0.98	0.02	0.28	0.28	0.08	95.60
α_3	0.1	0.09	0.01	0.30	0.30	0.09	96.70
Survival model of no-progression population							
β_{00}	-1	-1.01	0.01	0.26	0.23	0.07	89.01
β_{01}	1	0.96	0.04	0.25	0.25	0.06	97.80
β_{02}	0.2	0.21	0.01	0.28	0.24	0.08	93.40

Table 4.22: Simulation Results when $\sigma^2 = 2.25$ and $n = 400$ with $\beta_3 = 1$.

Disease progression model of progression population

	true	APE	bias	SD	ASD	MSE	CP
β_{10}	-0.5	-0.46	0.04	0.31	0.33	0.10	94.51
β_{11}	1	1.03	0.03	0.30	0.26	0.07	94.51
β_{12}	0	0.01	0.01	0.20	0.20	0.04	93.41

Gap time model of progression population

β_{20}	-0.3	-0.30	0.00	0.33	0.34	0.11	94.51
β_{21}	-0.5	-0.49	0.01	0.28	0.27	0.08	95.61
β_{22}	0.6	0.61	0.01	0.19	0.18	0.04	89.01
β_{23}	-0.5	-0.47	0.03	0.28	0.26	0.07	93.41
β_{24}	0.5	0.53	0.03	0.26	0.24	0.07	93.41
β_{25}	-0.4	-0.43	0.03	0.37	0.34	0.14	92.31

Table 4.23: Simulation Results when $\sigma^2 = 2.25$ and $n = 1000$ with $\beta_3 = 1$.

	true	APE	bias	SD	ASD	MSE	CP
σ^2	2.25	2.12	0.13	0.34	0.36	0.13	96.59
Susceptibility Model							
α_0	1.6	1.63	0.03	0.29	0.27	0.09	94.32
α_1	-1.8	-1.86	0.06	0.38	0.32	0.15	89.77
α_2	1	0.99	0.01	0.28	0.28	0.08	95.45
α_3	0.1	0.08	0.02	0.30	0.30	0.09	95.45
Survival model of no-progression population							
β_{00}	-1	-1.03	0.03	0.27	0.23	0.07	87.50
β_{01}	1	0.96	0.04	0.27	0.25	0.07	94.32
β_{02}	0.2	0.20	0.00	0.30	0.24	0.09	93.18

Table 4.24: Simulation Results when $\sigma^2 = 2.25$ and $n = 1000$ with $\beta_3 = 1$.

Disease progression model of progression population

	true	APE	bias	SD	ASD	MSE	CP
β_{10}	-0.5	-0.44	0.06	0.32	0.32	0.11	94.32
β_{11}	1	1.00	0.00	0.29	0.26	0.08	95.45
β_{12}	0	0.01	0.01	0.21	0.20	0.04	93.18

Gap time model of progression population

β_{20}	-0.3	-0.28	0.02	0.34	0.33	0.12	93.18
β_{21}	-0.5	-0.47	0.03	0.28	0.26	0.08	97.73
β_{22}	0.6	0.59	0.01	0.19	0.18	0.04	90.91
β_{23}	-0.5	-0.49	0.01	0.25	0.26	0.07	96.59
β_{24}	0.5	0.53	0.03	0.26	0.24	0.07	92.05
β_{25}	-0.4	-0.41	0.01	0.36	0.33	0.13	94.32

Chapter 5

Time-Varying Coefficient Cox Model for Access Change in Dialysis Study

5.1 Introduction

Dialysis treatment is necessary to sustain life for patients with end stage renal disease (ESRD). Dialysis treatment requires access to the bloodstream to deliver a high enough blood flow. Catheter (CATH) and arteriovenous (AV) are common methods of vascular access. The influence of using CATH and AV on the mortality risk is of great interest. Studies show that patients using CATH have a greater mortality risk than those using AV access (Astor, Eustace, Powe, Klag, Fink and Coresh 2005, Allon, Daugirdas, Depner, Greene, Ornt and Schwab 2006, Wystrychowski, Kitzler, Thijssen, Usvyat, Kotanko and Levin 2009, Lacson, Wang, Lazarus and Hakim 2009, Bradbury, Chen, Fumiss, Pisoni, Keen, Mapes and Krishnan 2009, Ng, Hung, Wu and Ko 2014). In practice, CATH, as a method of

vascular access in unprepared HD situations, usually serves patients only in the short term and patients may switch to AV as the method of long term vascular access. To evaluate the impact of time at access change on mortality risk may play an essential role to maximize the benefit from vascular access. However, previous studies have left this question unanswered. This chapter aims to investigate the association between time of access change and mortality risk through the Cox proportional hazard (Cox PH) model with time-varying coefficient, as well as to provide evidence for the hypothesis that patients with early access change from CATH to AV have lower mortality risk than those with late access change.

5.2 The Data

The data set consists of patients who were treated with HD at Fresenius medical care. The study lasted from January 2010 to December 2014. The demographic characteristics, *age*, *race*, *male*, *hispanic* and *height*, are recorded for all patients. Note *race* denotes white as 1 and all others as 0. The following 12 measures of baseline characteristics are taken: *diabetic*, pre-dialysis systolic blood pressure (*presbp*), pre-dialysis diastolic blood pressure (*predbp*), pre-dialysis weight (*preweight*), post-dialysis weight (*postweight*), the blood flow rate (QB), the dialysis fluid flow rate (QD), interdialytic weight gain (*idwg*), ultrafiltration rate (*ufr*), albumin (*album*), EPO use (*epo*), and neutrophil-to-lymphocyte ratio (*nlr*). Vas-

cular accesses are defined as catheter (CATH) and arteriovenous (AV), where AV includes patients with fistulas or grafts.

3835 patients who began HD with vascular access as CATH are included. A categorical variable *switch* is defined to be 1 when the access change is observed and 0 when no access change is observed during the study. The continuous variable *switchtime* denotes time at access change since the initiation of dialysis. The access change and censor rates are 57.57% and 74.97%, respectively. The baseline characteristic measures are taken as the measures in the first month.

For demographic characteristics, the average age is 65.97 with SD 15.34, the average height is 167.66 cm with SD 12.45, more patients are white (67.77% with SD 46.74), a slightly greater number of male patients than female ones (54.60% with SD 49.79), less than one sixth of patients are hispanic (13.06%, with SD 33.70) and more diabetic patients are enrolled (60.34%, 48.93). The descriptive statistics for baseline characteristics is summarized in Table 5.1 with two columns, mean and standard deviation (SD). We note that the variation of *epo* is very large in (SD: 5392.16). The primary outcome is the observed survival time with the unit as year (*eventtime*).

The overall goal of this study is to investigate the time-varying effect of access change. We apply the Cox PH model with time-varying coefficient to explore the relation between survival hazard and time at access change.

5.3 Time-Varying Coefficient Cox Model

Various methods have been proposed to fit a Cox PH model with varying co-effcient. While the association between the switch and failure event also varies with the time of switch, a Cox PH model with time-varying coefficient allows us to quantify this relation between switch and failure time through assuming that the coefficient of the time-dependent indicator of the switch event is a nonparametric function of time at switch.

Define T_D as the failure time, T_V as the time of access change and $U(t)$ as the access change event as time-dependent covariate, that is $U(t)$ equals to 1 as $t \geq T_V$ and 0 as $t < T_V$ for $T_D = t$. We consider the following Cox PH with varying coefficient

$$h_D(t|U(t), \mathbf{X}) = h_0(t)\exp\left[\beta(s)U(t) + \gamma\mathbf{X}\right], \tag{5.1}$$

where $h_0(t)$ denotes baseline hazard and $h_D(t)$ denotes hazard function for failure.

Table 5.1: Means and standard deviations of variables

	mean	*SD*
diabetic(%)	60.34	48.93
preweight (kg)	83.81	23.54
postweight (kg)	81.97	23.22
presbp (mmHg)	143.68	21.74
predbp (mmHg)	76.47	13.21
QB (mL/min)	367.63	42.90
QD (mL/min)	661.88	91.95
idwg (%)	2.93	1.35
ufr (%)	6.76	3.24
album (g/dl)	3.65	0.37
epo (units)	**5040.97**	**5392.16**
nlr (unitless)	6.01	5.53
eventtime	1.25	1.27
switchtime	0.19	0.34

5.4 Estimation

Perperoglou, le Cessie and van Houwelingen (2006) proposed a fast routine for fitting the Cox PH model with time-varying effects of the covariates, and Thomas and Reyes (2014) described the survival estimation in SAS and R for such a model. Nan, Lin, Lisabeth and Harlow (2005) proposed a Cox PH model with time-varying coefficient to investigate the association between age at a marker event, defined as a specific bleeding pattern change, and age at menopause, where both events are subject to censoring and their association varies with age at the marker event.

In this section, we will introduce the method to estimate the nonparametric function $\beta(s)$ using the regression spline method. Specifically, we approximate $\beta(s)$ using the natural cubic B-spline basis. Let K be the number of interior knots. The locations of knots are usually chosen such that an equal number of observed data values sits between any two adjacent knots, or knots are placed equally spaced through the study time.

The function $\beta(s)$ can be parameterized using $K+2$ natural cubic B-spline basis functions $B_j(s)$ for $j = 1, \ldots, K+2$ as

$$\beta(s) = \sum_{j=1}^{K+2} \theta_j B_j(s). \tag{5.2}$$

Substituting $\beta(s)$ by the approximation of B-spline in equation (5.2), we could rewrite model (5.1) as

$$h_{D_u}(t|U(t),\mathbf{X}) = h_0(t)\exp\left[\theta^{\mathbf{T}}\tilde{\mathbf{U}}(\mathbf{t}) + \gamma\mathbf{X}\right] \tag{5.3}$$

where $\theta = (\theta_1,\ldots,\theta_{K+2})^T$ and $\tilde{U}(t) = \{B_1(s)Z(t),\ldots,B_{K+2}(s)Z(t)\}^T$.

Now model (5.3) becomes a standard Cox proportional hazards model with time-dependent covariate vector $\tilde{U}(t)$ and the baseline covariate vector $\mathbf{X}$. Therefore, we can obtain the estimates of parameter (θ,γ) through partial likelihood estimators method. The no-parametric function $\beta(s)$ can be estimated by equation (5.2).

Standard methods may be applied to get estimates of $\beta(s)$ and γ when the knots are fixed. In the following we adopt the GCV method in Nan et al. (2005) to select the number of knots based on data.

Nan et al. (2005) considered the case when the cumulative baseline hazard function $H_0(t)$ is unknown and is estimated by the Breslow estimator,

$$\hat{H}_0(t) = \int_0^t \left[\sum_{i=1}^n I(Y_i \geq u)\exp\left[\hat{\beta}(S_i)Z_i(u) + \hat{\gamma}'X_i\right]\right]^{-1}\left[\sum_{i=1}^n dN_i(u)\right], \tag{5.4}$$

where $N_i(u)$ denotes the failure event counting process, $I(Y_i \geq u)$ denotes at risk process.

A series of Cox PH models as in (5.3) are fitted for a range of the number of interior knots K. For each choice of K, the cumulative baseline hazard function estimator $\hat{H}_0(t, K)$ is calculated. For a given number of knots K, the likelihood function of (θ, γ) is obtainable and can be maximized using an iterated reweighed least square algorithm. Denote the estimators of (θ, γ) at the lth iteration are $(\hat{\theta}_{(l)}, \hat{\gamma}_{(l)})$, the working weight w_i and the working dependent variable y_i for subject i are written as

$$
\begin{aligned}
w_i &= \frac{1}{2}\hat{H}_0(Y_i)\exp\left[\hat{\theta}'_{(l)}\tilde{U}_i(Y_i) + \hat{\gamma}'_{(l)}X_i\right], \\
y_i &= \hat{\theta}'_{(l)}\tilde{Z}_i(Y_i) + \hat{\gamma}'_{(l)}X_i + \triangle_i/(2w_i) - 1.
\end{aligned} \tag{5.5}
$$

where $\triangle_i$ denotes the censoring indicator. We calculates $(\hat{\theta}_{(l+1)}, \hat{\gamma}_{(l+1)})$ by minimizing $\sum_{i=1}^{n} \omega_i \left[y_i - \theta'\tilde{U}_i(Y_i) - \gamma' X_i\right]^2$. Denote the working dependent variable, the working weight matrix, and the predicted value vector at convergence by $\hat{\mathbf{y}} = (\hat{y}_1, \ldots, \hat{y}_n)', \hat{\mathbf{W}} = \text{diag}(\hat{w}_1, \ldots, \hat{w}_n)$, and $\hat{\mathbf{f}} = (\hat{f}_1, \ldots, \hat{f}_n)'$. Note that $\hat{\mathbf{f}}$ can be calculated as $\hat{\mathbf{f}} = \hat{\mathbf{X}}(\hat{\mathbf{X}}'\hat{\mathbf{W}}\hat{\mathbf{X}})^{-1}\hat{\mathbf{X}}\hat{\mathbf{W}}\hat{\mathbf{y}} = \hat{\mathbf{H}}\hat{\mathbf{y}}$, where $\hat{\mathbf{H}}$ is the linearized hat matrix. The GCV criterion as a function of knots K, is given by

$$\text{GCV}(K) = \frac{\sum_{i=1}^{n} w_i(\hat{y}_i - \hat{f}_i)^2}{(1-\hat{h})^2}, \tag{5.6}$$

where $\hat{h}$ is the average of the diagonal elements of $\hat{\mathbf{H}}$. We substitute different baseline hazard estimators for different K into (5.5) and (5.6). We then select K to minimize GCV(K), call it K_*, and obtain corresponding baseline hazard estimator $\hat{H}_0(t, K_*)$. Recalculating the GCV statistics in (5.6) using (5.5) by substituting the $\hat{H}_0(t, K_*)$ for each possible value of K and select a new K, as updated K_*, that minimize $GCV(K)$. Note a common $\hat{H}_0(t; K)$ is used to calculate GCV for different possible values of K. Above procedure is repeated until the chosen K_* at the current step is the same as the K_* at the previous step.

5.5 Analysis of the Dialysis Study

The Cox PH model with time-varying coefficient, including the time-dependent indicator of access change, *switch*, and all demographic and baseline characteristics, is fitted with the introduced main dataset of dialysis study to explore the time-varying effect of access change on mortality over time. Observed times at access change are used to generate the natural cubic B-spline basis function for estimating $\beta(s)$. Using a GCV algorithm, the obtained optimal number of interior knots is 5. The obtained estimates and the plot of the estimate of $\beta(s)$ for this

model are summarized in Table 5.2, and Figure 5.1.

Figure 5.1 describes the effect of access change on hazard throughout the study time. The X-axis stands for the time line for access change, while Y-axis stands for the hazard ratio of access change. The short blue segments at the bottom indicate the participating subjects. The red dashed line denotes reference, which takes value 1, as no impact on mortality risk. The solid black line is the estimated effect of access change using the B-spline basis, and the dashed navy lines are the 95% CI for the effect of access change. We observe that the 95% CI is well below the red reference line before 0.5 year. This indicates that the effect of access change before 0.5 year is significantly different from zero. Specifically, the access change at early stage has significant negative effect on mortality risk. The 95% CI for effect of access change after 0.5 year, however, contains one, implying that the impact on mortality risk is not significant. That is, the late access change has small impact on mortality risk.

The effects from eight demographic and baseline characteristics are found to be significant. The positive effects from two demographic and three baseline characteristics, *age*, *race*, *diabetic*, *idwg*, and *ufr*, on mortality risk appear to be significant. The negative effects from one demographic and two baseline characteristics, *hispanic*, *presbp* and *album*, on mortality risk appear to be significant.

Table 5.2: Estimates for Cox PH model with time-varying coefficient

Cox PH model	w/ time-varying coef.		
	EST	exp(EST)	p-value
age	**0.04 (0.00)**	**1.04**	**<0.001**
race	**0.42 (0.08)**	**1.53**	**<0.001**
male	-0.02 (0.08)	0.98	0.85
hispanic	**-0.34 (0.11)**	**0.71**	**0.00**
height (cm)	-0.00 (0.00)	0.92.	0.60
diabetic	**0.24 (0.07)**	**1.27**	**0.00**
preweight (kg)	-0.08 (0.04)	0.92	0.08
postweight (kg)	0.08 (0.05)	1.08	0.08
presbp (mmHg)	**-0.01 (0.00)**	**0.99**	**<0.001**
predbp (mmHg)	- 0.00 (0.00)	1.00	0.81
QB (mL/min)	-0.00 (0.00)	1.00	0.24
QD (mL/min)	-0.00 (0.00)	1.00	0.41
idwg (%)	**0.09 (0.04)**	**1.10**	**0.01**
ufr (%)	**0.04 (0.02)**	**1.04**	**0.01**
album (g/dl)	**-0.54 (0.10)**	**0.58**	**<0.001**
epo (units)	0.00 (0.00)	1.00	0.20
nlr (unitless)	0.01 (0.01)	1.01	0.18

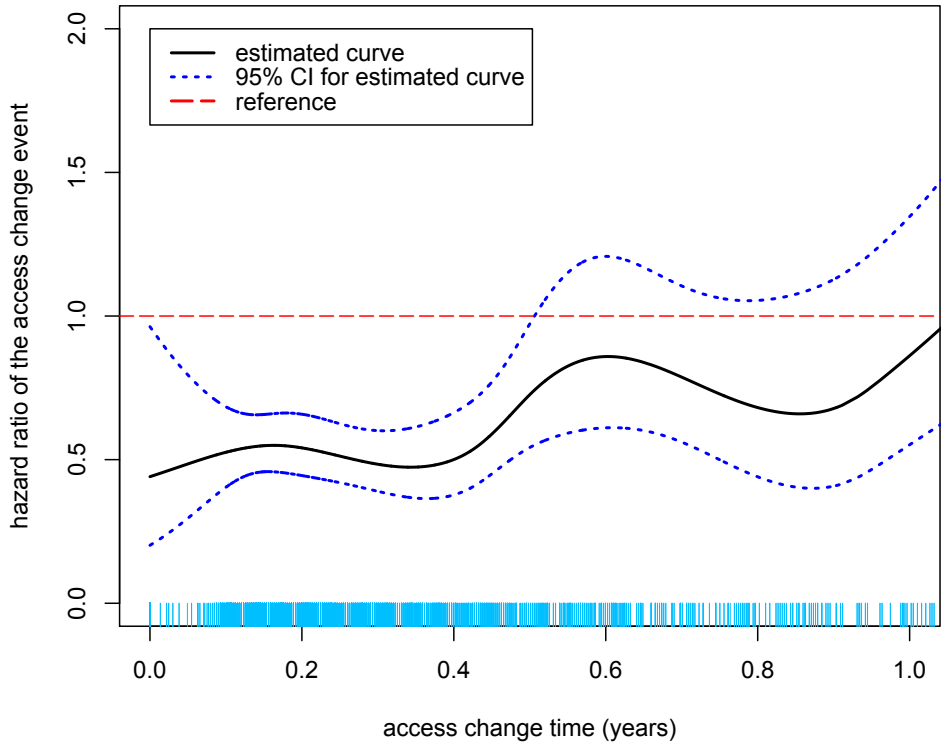

Figure 5.1: Hazard ratio of the effect of access change ($e^{\beta(s)}$): solid black line is the estimated effect of access change, and dashed navy lines are 95% CI for the effect of access change over the first year of the study.

Chapter 6

Accelerated Failure Time Model for Access Change in Dialysis Study

6.1 Introduction

An accelerated failure time (AFT) model as an extension of Robins and Tsiatis (1991)'s model is considered. For a patient who has access change, the latent event time is defined as an abstract quantity represents the patient's survival time that would have been if this patent had no access change. For a patient with access change, the latent event time is the same as his/her survival time. The effect of access change then can be estimated through the likelihood function. Given time at access change, Robins and Tsiatis (1991) estimated the effect of access change from a conditional likelihood function. While we model time at access change as a random variable, we estimate the effect of access change from a full likelihood

function.

6.2 Accelerated Failure Time Model with Latent Survival Time

We are using the following notations in this chapter,

(1) Indicator for access at the initiation of dialysis $R = \begin{cases} 0 & \text{CATH,} \\ 1 & \text{AV.} \end{cases}$

(2) Indicator for censoring to failure time $\delta = \begin{cases} 0 & \text{failure time censored,} \\ 1 & \text{failure time uncensored.} \end{cases}$

(3) Indicator for access change $V = \begin{cases} 0 & \text{no access change,} \\ 1 & \text{access change.} \end{cases}$

(4) T_0 denotes observed failure time for CATH, as the reference failure time.

(5) $S_0(t)$ denotes the survival function for failure time for patients with CATH.

(6) $f_0(t)$ denotes the probability density function for failure time for patients with CATH.

(7) T_1 denotes observed time at access change when initial access as CATH.

(8) $S_1(t)$ denotes the survival function for time at access change when initial access as CATH.

(9) $f_1(t)$ denotes the probability density function for time at access change when initial access as CATH.

(10) T_V denotes time at access change.

(11) $S_V(t)$ denotes the survival function for time at access change.

(12) $f_V(t)$ denotes the probability density function for time at access change.

(13) T_D denotes failure time.

(14) $S_D(t)$ denotes the survival function for failure time.

(15) $f_D(t)$ denotes the probability density function for failure time.

(16) T_L denote failure time when the access change is present.

(14) $S_L(t)$ denotes the survival function for failure time when the access change is present.

(15) $f_L(t)$ denotes the probability density function for failure time when the access change is present.

(16) $\exp[\alpha_1]$ denotes the effect from AV on failure time when no access change occurs.

(17) $\exp[\alpha_2]$ denotes the effect from AV on time at access change.

(18) $\exp[\alpha_3]$ denotes effect of access change from CATH to AV.

(19) $\exp[\alpha_4]$ denotes effect of access change from AV to CATH.

(20) $\mathbf{X}$ represents baseline covariates.

(21) $\mathbf{Z}$ represents covariates collected at access change as prognostic factors for the decision to access change.

(22) $\exp[\beta_1]$ denotes the effect from baseline covariates on failure time when no access change occurs.

(23) $\exp[\beta_2]$ denotes the effect from baseline covariates on time of access change before access change occurs.

(24) $\exp[\beta_3]$ denotes the effect from baseline covariates on time at access change to death.

(25) $\exp[\beta_4]$ denotes effect from prognostic factors on time at access change to death.

The observed event times and times at access change are assumed to be

$$T_D \stackrel{d}{=} \exp[\alpha_1 R + \beta_1 \mathbf{X}]\, T_0, \tag{6.1}$$

$$T_V \stackrel{d}{=} \exp[\alpha_2 R + \beta_2 \mathbf{X}]\, T_1. \tag{6.2}$$

Consider the situation where patients are allowed to have access change, using Robins and Tsiatis (1991)'s model, for a patient with access change at time T_V,

the latent event time T_D is defined as an abstract quantity represents the patient's survival time that would have been if this patent had no access change while the actual event time T_L is defined as the survival time when the access change is present. For a patient without access change, the latent event time is the same as his/her survival time. A simplified graph describes the relation of T_L and T_D in Figure 6.1.

For patients with access change from CATH to AV or AV to CATH, as $T_D \geq T_V$, the latent and observed event times can be related as follows

$$T_L \stackrel{d}{=} T_V + \exp\left[\alpha_3(1-R) + \alpha_4 R + \beta_3 \mathbf{X} + \beta_4 \mathbf{Z}\right](T_D - T_V). \tag{6.3}$$

6.3 Estimation

In this section, we construct the full likelihood by the following steps: (1) categorize subjects into four group according to their experienced events, (2) derive the likelihood function contributed by each group and (3) construct the full likelihood function with the contributed part by each group.

Let Y denote the observed failure time, and S denote the observed time at access change. $V = 0$ indicates the failure occurs before access change, while

$V = 1$ indicates the failure occurs after access change. $\delta = 1$ denotes uncensored and $\delta = 0$ denotes censored. The observed data can be divided into four groups of observations:

Group 1. Subjects are observed to fail at time Y without access change event occurring. These subjects have $V = 0, T_D = Y$ and $\delta = 1$.The contribution to the likelihood function is

$$f_D(Y|R,\mathbf{X})S_V(Y|R,\mathbf{X}) = \exp\left[-\alpha_1 R - \beta_1 \mathbf{X}\right] f_0\big(\exp\left[-\alpha_1 R - \beta_1 \mathbf{X}\right] Y\big) \\ S_1\big(\exp\left[-\alpha_2 R - \beta_2 \mathbf{X}\right] Y\big). \tag{6.4}$$

Group 2. The subjects are observed with time at access change S and fail at time Y. The subjects have $V = 1$, $\delta = 1$, $T_L = Y$, $T_V = S$. The contribution to the likelihood function is

$$\begin{aligned} &f_{L,V}(Y,S|R,\mathbf{X},\mathbf{Z}) \\ &= f_V(S|R,\mathbf{X}) f_L(Y|T_V = S,R,\mathbf{X},\mathbf{Z}) \\ &= \exp\left[-\alpha_2 R - \beta_2 \mathbf{X}\right] f_1\big(\exp\left[-\alpha_2 R - \beta_2 \mathbf{X}\right] S\big) \\ &\times \exp\left[-\alpha_1 R - \beta_1 \mathbf{X} - \alpha_3(1-R) - \alpha_4 R - \beta_3 \mathbf{X} - \beta_4 \mathbf{Z}\right] \\ &\times f_0\Big(\exp\left[-\alpha_1 R - \beta_1 \mathbf{X}\right]\left(S + \exp\left[-\alpha_3(1-R) - \alpha_4 R - \beta_3 \mathbf{X} - \beta_4 \mathbf{Z}\right](Y-S)\right)\Big). \end{aligned} \tag{6.5}$$

Group 3. The subjects are observed with time at access change S and censored at time Y. These subjects have $V = 1$, $\delta = 0$, $T_L \geq Y$, and $T_V = S$. The contribution to the likelihood function is

$$
\begin{aligned}
&P(T_L \geq Y, T_V = S|R, \mathbf{X}, \mathbf{Z}) \\
&= f_V(S|R, \mathbf{X}) P(T_L \geq Y|T_V = S, R, \mathbf{X}, \mathbf{Z}) \\
&= f_V(S|R, \mathbf{X}) S_D(S + \exp\left[-\alpha_3(1-R) - \alpha_4 R - \beta_3 \mathbf{X} - \beta_4 \mathbf{Z}\right](Y - S)|R, \mathbf{X}, \mathbf{Z}) \\
&= \exp\left[-\alpha_2 R - \beta_2 \mathbf{X}\right] f_1\Big(\exp\left[-\alpha_2 R - \beta_2 \mathbf{X}\right] S\Big) \\
&\times S_0\bigg(\exp\left[-\alpha_1 R - \beta_1 \mathbf{X}\right]\Big(S + \exp\left[-\alpha_3(1-R) - \alpha_4 R - \beta_3 \mathbf{X} - \beta_4 \mathbf{Z}\right](Y - S)\Big)\bigg).
\end{aligned}
\tag{6.6}
$$

Group 4. Subjects are only observed to be censored at time Y without access change and death. These subjects have $V = 0$, $\delta = 0$ and $T_D > Y$; or, they may have $V = 1$, $\delta = 0$ and $T_V > Y$. These subjects may also have $V = 1$, $\delta = 0$ and $T_L > T_V > Y$, The contribution to the likelihood function is

$$
\begin{aligned}
&P(T_D \geq Y|R, \mathbf{X}) P(T_V \geq Y|R, \mathbf{X}) \\
&= S_D(Y|R, \mathbf{X}) S_V(Y|R, \mathbf{X}) \\
&= S_0\Big(\exp\left[-\alpha_1 R - \beta_1 \mathbf{X}\right] Y\Big) S_1\Big(\exp\left[-\alpha_2 R - \beta_2 \mathbf{X}\right] Y\Big).
\end{aligned}
\tag{6.7}
$$

Based on above four categorizations of group, the full log likelihood is derived as below:

$$\log L(\cdot|\mathbf{Y},\mathbf{S}) = \delta(1-V)(\log f_D(Y|R,\mathbf{X}) + \log S_V(Y|R,\mathbf{X}))$$
$$+ \delta V(\log f_V(S|R,\mathbf{X}) + \log f_L(Y|T_V = S,R,\mathbf{X},\mathbf{Z})$$
$$+ (1-\delta)V(\log f_V(S|R,\mathbf{X}) + \log P(T_L \geq Y|T_V = S,R,\mathbf{X},\mathbf{Z}))$$
$$+ (1-\delta)P(V=1)(\log S_D(Y|R,\mathbf{X}) + \log S_V(Y|R,\mathbf{X}))$$
$$+ (1-\delta)P(V=0)(\log S_D(Y|R,\mathbf{X}) + \log S_V(Y|R,\mathbf{X})). \qquad (6.8)$$

6.4 Simulation Study

We set $S_0(t) = \exp[-t]$, $S_1(t) = \exp[-t]$. Suppose $\alpha_1 = 1$, $\alpha_2 = 1.5$, $\beta_1 = 1$, $\beta_2 = 1.5$, $\beta_3 = 0.2$, $\beta_4 = 0.3$, $\alpha_3 = 5$, and $\alpha_4 = -1$. The initial access R is assigned to be 0 for the first half of the subjects and 1 for the other half. The failure time T_D and the time at access change T_V are generated using equation (6.2) and (6.3). We assign the indicator for access change V to be 0 when $T_D < T_V$, and to be 1 when $T_D > T_V$. For patients with $V = 1$, using equation 6.3, we further generated T_L with obtained T_D and T_V. The baseline covariate $\mathbf{X} \overset{\text{i.i.d.}}{\sim} \text{Ber}(0.7)$ and the prognostic covariate $\mathbf{Z} \sim \text{Uni}(0,1)$ Assume censoring time assumes to be $C \sim \text{Uniform}[0,72]$. The censoring indicator δ takes 1 when $T_L \leq C$ and 0 when $T_L > C$. Denote $Y = \min(T_L, C)$ as the observed failure time and observed time at access change $S = T_V$ for $V = 1$. The average access change and censoring rate for 100 sets of data with sample size as 500 and 1000 are (35.79%, 29.18%) and

(35.91%, 29.35%), respectively.

From 100 simulated datasets, the obtained MSE for the majority average of estimates are below 0.5 except for the estimate for α_3 of sample size 500. The slightly elevated MSE of α_3 is attributed to the increase SD between the estimates from 100 simulated dataset. All CP are close to 95% for both sample size 500 and 1000. We conclude that the estimation method performed well.

Table 6.1: Simulation result with access change rate as 35.79% and censoring rate as 29.18% and $n = 500$.

	$n =$	500				
	True	APE	Bias	ASD	SD	MSE	CP
α_1	1	0.99	0.01	0.10	0.11	0.01	93%
α_2	1.5	1.48	0.02	0.13	0.15	0.02	93%
α_3	5	5.32	0.32	7.27	1.91	3.75	98%
α_4	-1	-1.03	0.03	0.34	0.36	0.13	95%
β_1	1	1.02	0.02	0.09	0.09	0.01	94%
β_2	1.5	1.47	0.03	0.11	0.11	0.01	94%
β_3	2.5	2.53	0.03	0.34	0.37	0.14	97%
β_4	1.5	1.48	0.02	0.53	0.49	0.24	97%

6.5 Description of Dataset

Among total of 14468 patients, 6621 patient began HD with vascular access as CATH, with 3437 of them switched to AV, and 7245 patients began HD with vascular access as AV, with 1195 of them switched to CATH. A categorical variable *switch* is defined to be 1 when the access change is observed and 0 when no access change is observed during the study. The continuous variable *switchtime* denotes time at access change. The average access change and censor rates are 33.03% and 74.18%, respectively. The access change rate is 51.91% for patient

Table 6.2: Simulation result with access change rate as 35.91% and censoring rate as 29.35% and $n = 1000$.

	$n =$	1000				
	True	APE	Bias	ASD	SD	MSE	CP
α_1	1	1.00	0.00	0.07	0.08	0.01	89%
α_2	1.5	1.47	0.03	0.09	0.08	0.01	95%
α_3	5	5.19	0.19	0.40	0.43	0.22	97%
α_4	-1	-1.01	0.01	0.24	0.23	0.05	95%
β_1	1	1.02	0.02	0.06	0.07	0.001	93%
β_2	1.5	1.48	0.02	0.08	0.07	0.01	99%
β_3	2.5	2.49	0.01	0.24	0.25	0.06	93%
β_4	1.5	1.46	0.04	0.38	0.41	0.17	93%

with initial access as CATH and is 16.49% for patients with initial access as AV. The baseline characteristic measures are taken as the measures in the first month.

The descriptive statistics for demographic, baseline and prognostic characteristics are summarized in Table 6.3 and 6.4. The statistics are presented in the format of mean (SD). The primary outcome is the observed survival time with year as unit (*eventtime*). While more patients of *race*:white than all other race (63.60%) are enrolled, there are a lightly greater number of male than female patients (56.41%). Less than one sixth of patients are hispanic (14.07%). More diabetic patients are enrolled (60.62%). Note that the variation of *epo* is very large in both datasets (SD: 5419.62).

The overall goal of this study is to investigate the effect of access change from CATH to AV and from AV to CATH on survival time, assuming time at access change, T_V, as a random variable.

6.6 Analysis with AFT model for Dialysis Dataset

The AFT model is fitted with the dialysis dataset to explore the effect of access change from CATH to AV and from AV to CATH. Estimates are summarized in Table 6.5 - 6.10.

Table 6.3: Descriptive statistics of demographic characteristics

	mean	SD
age	66.09	15.01
race(%)	63.60	48.12
male(%)	56.41	49.59
hispanic(%)	14.07	34.77
height (cm)	167.72	12.47

The positive effects of AV on survival and time at access change are found to be significant. While the positive effect of access change from CATH to AV appear to be significant, the effect of access change from AV to CATH is found to be insignificant. Specifically, the effect of AV on survival time for patients without access change is to prolong the time by the factor 2.14 with 95% confidence interval $(1.95, 2.32)$, while the effect of AV on time at access change is to prolong the time by the factor 11.70 with 95% confidence interval $(10.91, 12.55)$. The effect of access change from CATH to AV on survival time is to prolong the time by the factor 5.58 with 95% confidence interval $(1.08, 28.79)$. Note that effects of staying with CATH on survival time and time at access change are defined as reference.

In Tables 6.7 and 6.8, two negative and one positive demographic effects, *age*, *race* and *hispanic*, on survival are found to be significant for patients without

Table 6.4: Descriptive statistics of baseline and prognostic characteristics

	baseline		prognostic	
	mean	SD	mean	SD
diabetic(%)	60.62	48.86		
preweight (kg)	83.68	23.39	83.41	24.10
postweight (kg)	81.60	23.01	81.20	22.99
presbp (mmHg)	146.06	21.73	147.26	21.22
predbp (mmHg)	76.40	12.93	77.56	12.95
QB (mL/min)	380.14	57.74		
QD (mL/min)	667.79	93.37		
idwg (%)	2.66	1.39		
ufr (%)	7.63	3.57	7.73	3.19
album (g/dl)	3.59	0.53	3.66	0.46
epo (units)	4972.08	5419.62		
nlr (unitless)	5.40	5.02	4.53	4.04
eventtime	1.55	1.39		
switchtime	0.18	0.47		

Table 6.5: Estimates for effect from AV for the dialysis dataset

	T_D		T_V	
	EST	p-value	EST	p-value
AV	**0.76 (0.04)**	**<0.001**	**2.46 (0.04)**	**<0.001**

access change. Minor positive effect of unit increase in *age* and negative effect of being white and an unit increase in *height* on time at access change are also found to be significant.

The effects on survival time for patients without access change and the effects on time at access change from baseline characteristics are summarized in Table 6.9. Six negative effects, *diabetic*, *postweight*, *idwg*, *ufr*, *epo*, and *nlr*, along with three positive effects, *preweight*, *presbp* and *album* on survival time for patients without access change are found to be significant. Three negative effects, *diabetic*, *QD*, and *nlr*, along with three positive effects, *QB*, *idwg* and *album* on time at access change are found to be significant.

The effects for baseline and prognostic characteristics on survival for patients with access change are summarized in Table 6.10. Two negative effects, *presbp* and *album*, along with one positive effects, *nlr* of baseline characteristics on survival for patients with access change are found to be significant.

Table 6.6: Estimates for effect of access change for the dialysis dataset

	T_L	
	EST	p-value
CATH to AV	**1.72 (0.84)**	**0.04**
AV to CATH	0.86 (0.85)	0.31

Table 6.7: Estimates for demographic characteristics of the failure time without access change, T_D, and time at access change, T_V

	T_D		T_V	
	EST	p-value	EST	p-value
age	**-0.03 ($\approx$ 0.00)**	**<0.001**	**$\approx$ 0.00 ($\approx$ 0.00)**	**<0.001**
race	**-0.33 (0.04)**	**<0.001**	**-0.08 (0.03)**	**0.02**
male	-0.08 (0.04)	0.07	-0.06 (0.03)	0.06
hispanic	**0.29 (0.06)**	**<0.001**	0.02 (0.04)	0.60
height (cm)	0.00 ($\approx$ 0.00)	0.08	**-0.00 ($\approx$ 0.00)**	**<0.001**

Table 6.8: Estimates for demographic characteristics of the failure time with access change, T_L

	T_L	
	EST	p-value
age	-0.00 ($\approx$ 0.00)	0.17
race	-0.03 (0.09)	0.75
male	-0.01 (0.09)	0.90
hispanic	0.01 (0.12)	0.91
height (cm)	-0.00 ($\approx$ (0.00)	0.29

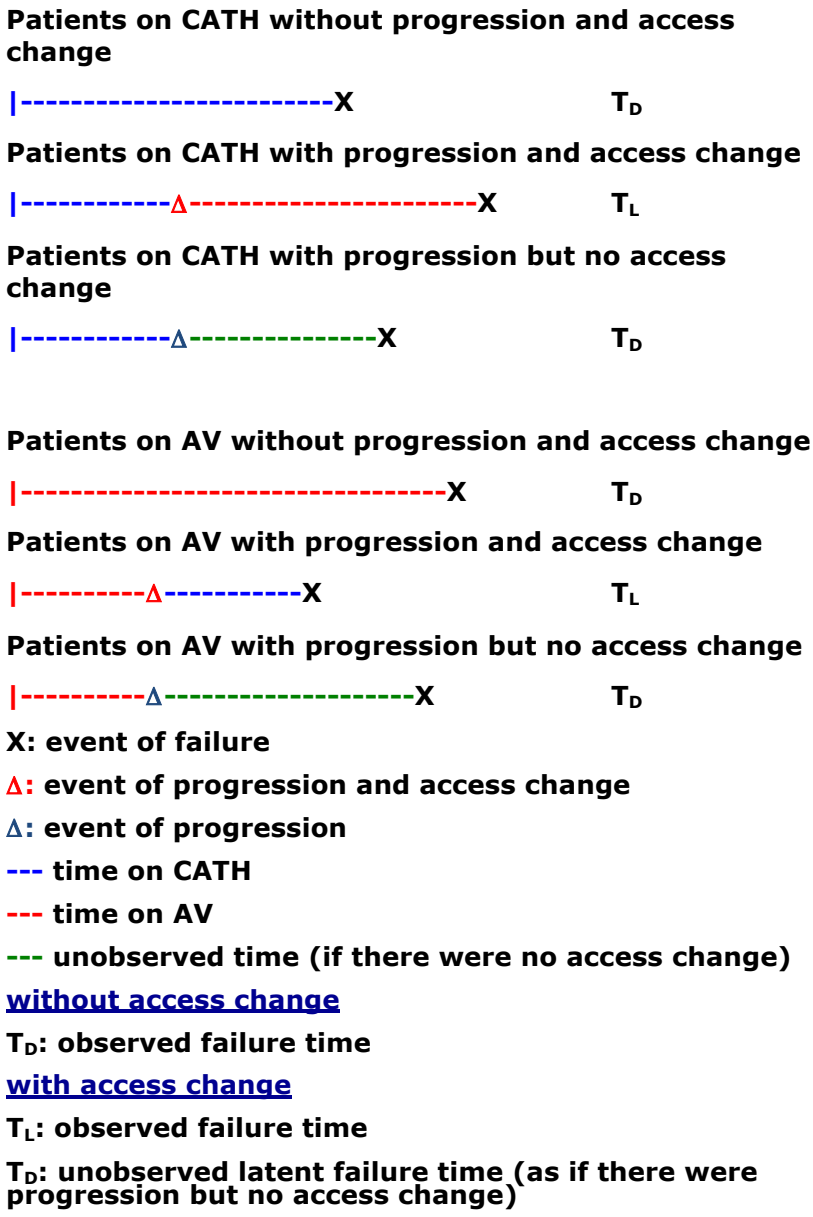

Figure 6.1: Relation in graph between T_D and T_L with/without access change.

Table 6.9: Estimates for baseline characteristics of failure time without access change (T_D) and time at access change (T_V)

	T_D		T_V	
	EST	p-value	EST	p-value
diabetic	**-0.15 (0.04)**	**<0.001**	**-0.15 (0.03)**	**<0.001**
preweight (kg)	**0.06 (0.03)**	**0.01**	-0.00 (0.02)	0.88
postweight (kg)	**-0.06 (0.03)**	**0.02**	0.00 (0.02)	0.91
presbp (mmHg)	**0.01 (≈ 0.00)**	**<0.001**	-0.00 (≈ 0.00)	0.30
predbp (mmHg)	0.00 (≈ 0.00)	0.61	-0.00 (≈ 0.00)	0.86
QB (mL/min)	-0.00 (≈ 0.00)	0.15	**0.00 (≈ 0.00)**	**0.01**
QD (mL/min)	-0.00 (≈ 0.00)	0.77	**-0.00 (≈ 0.00)**	**0.01**
idwg (%)	**-0.05 (0.02)**	**0.04**	**0.07 (0.02)**	**<0.001**
ufr (%)	**-0.05 (0.01)**	**<0.001**	0.01 (0.01)	0.12
album (g/dl)	**0.68 (0.04)**	<0.001	**0.15 (0.03)**	**<0.001**
epo (units)	**-0.00 (≈0.00)**	**<0.001**	0.00 (≈ 0.00)	0.25
nlr (unitless)	**-0.23 (≈ 0.00)**	**<0.001**	**-0.01 (≈ 0.00)**	**0.01**

Table 6.10: Estimates for baseline and prognostic characteristics of failure time when access change is present (T_L)

	T_L			
	baseline		prognostic	
	EST	p-value	EST	p-value
diabetic	-0.08 (0.09)	0.35		
preweight (kg)	-0.07 (0.05)	0.10	-0.05 (0.05)	0.24
postweight (kg)	0.06 (0.05)	0.20	0.01 (0.05)	0.15
presbp (mmHg)	**-0.01 (≈ 0.00)**	**<0.001**	**0.02 (≈ 0.00)**	**<0.001**
predbp (mmHg)	0.01 (0.01)	0.16	**-0.02 (0.01)**	**<0.001**
QB (mL/min)	0.00 (≈ 0.00)	0.42		
QD (mL/min)	-0.00 (≈ 0.00)	0.84		
idwg (%)	0.06 (0.05)	0.25		
ufr (%)	0.03 (0.02)	0.08	-0.01 (0.02)	0.82
album (g/dl)	**-0.75 (0.09)**	**<0.001**	**0.91 (0.08)**	**<0.001**
epo (units)	0.00 (≈ 0.00)	0.68		
nlr (unitless)	**0.03 (0.01)**	**<0.001**	**-0.03 (0.01)**	**<0.001**

Appendix A

SAS code

A.1 SAS code of fitting AFT model

```
libname out '/home/grad/chu/';

data out.simdata1n;
        infile "simdata1_expnorm_wonew.txt"; *you could change this
        to your own directory where the data is stored;
        input id$ R X1 X2 Z V W Y Group;

* R randomization regime;
* Group as 4 categories, indicating switching and censoring status;
* subject id
* V as switching status;
* X1 and X2 are baseline covariates;
* Z covariate collected at disease progression;
* Y observed failure time;
* W progression time;
run;

proc print data=out.simdata1n;
run;


***NLMIXED procedure: The integration method is specified to
take default method, adaptive Gauss-Hermite quadrature
(\textbf{GAUSS}).
The optimization is set to be default algorithm,
```

```
Quasi-Newton Optimization (\textbf{QUANEW}).
The update formula is specified to be the dual Davidon, Fletcher,
and Powell (DFP) update of the Cholesky factor of
the Hessian matrix (\textbf{DDFP}).
The maximum number of quadrature points (\textbf{QMAX})
before the adaptive search is aborted sets to be 100;

proc nlmixed data=out.simdata1n qmax=100 itdetails upd=ddfp;
***specify the parameters with bounds;
**** set initial values;
parms alpha0=1.6 alpha1=-1.8 alpha2=1 alpha3=0.1
      beta00=-1 beta01=1 beta02=0.2
      beta10=-0.5 beta11=1 beta12=0
      beta20=-0.3 beta21=-0.5 beta22=0.6 beta23=-0.5
      beta24=0.5 beta25=-0.4;
**assume the linear combination part of each accelerated factor;
 lin_e0 = beta00*R + beta01*X1 + beta02*X2;
 lin_e1 = beta10*R + beta11*X1 + beta12*X2;
 lin_e2 = beta20*R + beta21 *V* (1-R)+ beta22*W+ beta23*X1
     + beta24*X2+beta25*Z;


***the survival functions of $T_D$, $T_U$ and $T_G$ are written
based on equations in chapter 3.;
Correspondent hazard functions are obtained by
 the general formula of survival function;
***accelerated factor for T_D (time to death for non progression group),
T_U(time to progression) and T_G(time from progression to death);


  e_d = exp(lin_e0);
  e_u = exp(lin_e1);
  e_g = exp(lin_e2);
***logit model for probability of progression,
pu_1 as probability of progression and
pu_0 as probability of non-progression;

 pu_1 = 1/(1+exp(-(alpha0+alpha1*R+alpha2*X1+alpha3*X2)));
 pu_0 = 1 - 1/(1+exp(-(alpha0+alpha1*R+alpha2*X1+alpha3*X2)));
***survival function using formula;
*survival function of T_D;
                        S_d = exp(-Y * e_d);
```

```
*survival function of T_U, for the patients belong to group 2 and 3;
 S_u23 = exp(-W * e_u);
*survival function of T_U, for the patients belong to group 4;

S_u4 = exp(-Y * e_u);
*survival function of T_G;
S_g = exp(-(Y-W) *e_g);


***correspondent hazard function;

  h_d =  e_d;
  h_u =  e_u;
  h_g =  e_g;

***specify the contribution to the log likelihood function of subject
 who is categorized to belong to Group 1, 2, 3 and 4, respectively
***loglikelihood from 4 group of subjects;
* loglik for group 1(U=0, delta=1);
if (Group=1) then loglik = log (h_d) + log(S_d) + log( pu_0);

* loglik for group 2(U=1, delta=1);
if (Group=2) then loglik = log (h_u) + log(h_g) + log( S_u23)
    + log(S_g) + log( pu_1);
* loglik for group 3(U=1, delta=0);
if (Group=3) then loglik = log (h_u) + log( S_u23)
    + log( S_g) + log(pu_1);
* loglik for group 4(U=1/U=0, delta=1);
if (Group=4) then loglik = log (S_d *pu_0 + S_u4 * pu_1);

*** the full likelihood function;
*** A model statement is then specified
using the option general (loglik)
to describe the distribution of the response variable.
No random statement is specified for this case;
     model Y ~ general(loglik);
     ods output ParameterEstimates=out.est_wo FitStatistics=out.fit;

     run;
```

A.2 SAS code of fitting AFT model with a frailty term

```
libname out '/home/grad/chu/';

data out.simdata1n;
        infile "simdata1_expnorm_wnew.txt"; * you should change this
         to your own directory where the data is stored;
        input id$ R X1 X2 Z V W Y Group;

* R randomization regime;
* Group as 4 categories, indicating switching and censoring status;
* subject id;
* V as switching status;
* X1 and X2 are baseline covariates;
* Z covariate collected at disease progression;
* Y observed failure time;
* W progression time;
run;

proc print data=out.simdata1n;
run;



***NLMIXED procedure: The integration method is specified to
take default method, adaptive Gauss-Hermite quadrature
(\textbf{GAUSS}).
The optimization is set to be default algorithm,
Quasi-Newton Optimization (\textbf{QUANEW}).
The update formula is specified to be the dual Davidon, Fletcher,
and Powell (DFP) update of the Cholesky factor of
the Hessian matrix (\textbf{DDFP}).
The maximum number of quadrature points (\textbf{QMAX})
before the adaptive search is aborted sets to be 100;

proc nlmixed data=out.simdata1n qmax=100 itdetails upd=ddfp;
***specify the parameters with bounds;
**** set initial values;
parms sigma2=1 alpha0=1.6 alpha1=-1.8 alpha2=1 alpha3=0.1
```

```
          beta00=-1 beta01=1 beta02=0.2
          beta10=-0.5 beta11=1 beta12=0
          beta20=-0.3 beta21=-0.5 beta22=0.6 beta23=-0.5
          beta24=0.5 beta25=-0.4;

**assume the linear combination part of accelerated factor;
 lin_e0 = beta00*R + beta01*X1 + beta02*X2;
 lin_e1 = beta10*R + beta11*X1 + beta12*X2;
 lin_e2 = beta20*R + beta21 *V* (1-R)+ beta22*W
     + beta23*X1+ beta24*X2+beta25*Z;


***the survival functions of $T_D$, $T_U$ and $T_G$ are written
based on equations in chapter 3.
Correspondent hazard functions are obtained
by the general formula of survival function
***accelerated factor for T_D (time to death for non progression group),
T_U(time to progression) and T_G(time from progression to death);
***omega stands for frailty;


  e_d =  exp(lin_e0);
  e_u =  exp(lin_e1 + omega);
  e_g =  exp(lin_e2 + omega);

***logit model for probability of progression,
 pu_1 as probability of progression and
 pu_0 as probability of non-progression;

 pu_1 = 1/(1+exp(-(alpha0+alpha1*R+alpha2*X1+alpha3*X2)));
 pu_0 = 1 - 1/(1+exp(-(alpha0+alpha1*R+alpha2*X1+alpha3*X2)));
***survival function using formula;
*survival function of T_D;
                          S_d = exp(-Y * e_d);
*survival function of T_U, for the patients belong to group 2 and 3;
        S_u23 = exp(-W * e_u);
*survival function of T_U, for the patients belong to group 4;
  S_u4 = exp(-Y * e_u);
*survival function of T_G;
S_g = exp(-(Y-W) *e_g);
```

```
***correspondent hazard function;


  h_d =  e_d;
  h_u =  e_u;
  h_g =  e_g;

***specify the contribution to the log likelihood function of subject
who is categorized to belong to Group 1, 2, 3 and 4, respectively
***loglikelihood from 4 group of subjects;
* loglik for group 1(U=0, delta=1);
if (Group=1) then loglik = log (h_d) + log(S_d) + log( pu_0);
* loglik for group 2(U=1, delta=1);
if (Group=2) then loglik = log (h_u) + log(h_g) + log( S_u23)
      + log(S_g) + log( pu_1);
* loglik for group 3(U=1, delta=0);
if (Group=3) then loglik = log (h_u) + log( S_u23)
    + log( S_g) + log(pu_1);
* loglik for group 4(U=1/U=0, delta=1);
if (Group=4) then loglik = log (S_d *pu_0 + S_u4 * pu_1);


*** the full likelihood function;

model Y ~ general(loglik);
*** random statement is used to define normal random effects,
as frailty term, with subjects specified
 using the keyword \textbf{subject} in this statement.
This will result in a longormal component.
*** specify the distribution of frailty as normal
and the random factor assumed to be the variation among subjects;
     random omega ~ normal(0,sigma2) subject=id;
     ods output ParameterEstimates=out.est_w FitStatistics=out.fit;

        run;
```

Bibliography

Allon, M., Daugirdas, J., Depner, T., Greene, T. A., Ornt, D. and Schwab, S. J. (2006). Effect of change in vascular access on patient mortality in hemodialysis patients, *Am J Kidney Dis* **47(3)**: 469–477.

Astor, B. C., Eustace, J. A., Powe, N., Klag, M. J., Fink, N. E. and Coresh, J. (2005). Type of vascular access and survival among incident hemodialysis patients: The choices for healthy outcomes in caring for ESRD (CHOICE study, *J Am Soc Nephrol* **16(5)**: 1449–1455.

Bradbury, B. D., Chen, F., Fumiss, A., Pisoni, R. L., Keen, M., Mapes, D. and Krishnan, M. (2009). Conversion of vascular access type among incident hemodialysis patients: Description and association with mortality, *Am J Kidney Dis* **53(5)**: 804–814.

Branson, M. and Whitehead, J. (2002). Estimating a treatment effect in survival studies in which patients switch treatment, *Statistics in Medicine* **15: 21(17)**: 2449–63.

Collett, D. (2003). *Modelling Survival Data in Medical Research*, Chapman and Hall/CRC.

Fergusson, D., Aaron, S. D., Guyatt, G. and Herbert, P. (2002). Post-randomisation exclusions: the intention o treat principle and excluding patients from analysis, *British Medical Journal* **325**: 652–4.

Lacson, E. J., Wang, W., Lazarus, J. M. and Hakim, R. M. (2009). Change in vascular access and mortality in maintenance hemodialysis patients, *Am J Kidney Dis* **54(5)**: 912–21.

Law, M. G. and Kaldor, J. M. (1996). Survival analyses of randomized clinical trials adjusted for patients who switch treatments, *Statistics in Medicine* **15(19)**: 2069–76.

Morden, J. P., Lambert, P. C., Latimer, N., Abrams, K. R. and Wailoo, A. J. (2011). Assessing methods for dealing with treatment switching in randomized controlled trials, *BMC Medical Research Methodology* **11:4**: 1–20.

Nan, B., Lin, X., Lisabeth, L. D. and Harlow, S. D. (2005). A varying-coefficient cox model for the effect of age at a marker event on age at menopause, *Biometrics* **61(2)**: 576–63.

Ng, Y., Hung, Y., Wu, S. and Ko, P. (2014). Characteristics and 3-year mortality and infection rates among incident hemodialysis patients with a permanent catheter undergoing a first vascular access conversion, *Clin Exp Nephrol* **18(2)**: 329–38.

Perperoglou, A., le Cessie, S. and van Houwelingen, H. C. (2006). A fast routine for fitting cox models with time varying effects of the covariates, *Biomedicine* **81(2)**: 154–61.

Robins, J. M. and Tsiatis, A. A. (1991). Correcting for non-compliance in randomized trials using rank preserving structural failure time models, *Communications in Statistics - Theory and Methods* **20(8)**: 2609–31.

Shao, J., Chang, M. and Chow, S. C. (2005). Statistical inference for cancer trails with treatment switching, *Statistics in Medicine* **24(12)**: 1783–90.

Thomas, L. and Reyes, E. M. (2014). Tutorial: Survival estimation for cox regression models with time-varying coefficients using SAS and R, *Journal of Statistical Software* **61(01)**: 1–23.

Walker, A. S., White, I. R. and Babiker, A. G. (2004). Parametric randomization-based methods for correcting for treatment changes in the assessment of the causal effect of treatment, *Statistics in Medicine* **23(4)**: 571–590.

Wystrychowski, G., Kitzler, T. M., Thijssen, S., Usvyat, L., Kotanko, P. and Levin, N. W. (2009). Impact of switch of vascular access type on key clinical and laboratory parameters in chronic hemodialysis patients, *Nephrol Dial Transplant* **24(7)**: 2194–2200.

Zeng, D., Chen, Q., Chen, M. H., Ibrahim, J. G. and Amgen Research Group (2012). Estimating treatment effects with treatment switching via semi competing risks models: an application to a colorectal cancer study, *Biometrika* **99(1)**: 167–184.